DTIC FILE COPY

SECURITY CLASSIFICATION OF THIS PAGE

REPORT DOCUMENTATION PAGE

Form Approved OMB No. 0704-0188

1a. REPORT SECURITY CLASSIFICATION: UNCLASSIFIED	1b. RESTRICTIVE MARKINGS: NONE
2a. SECURITY CLASSIFICATION AUTHORITY: AD-A217 427	3. DISTRIBUTION/AVAILABILITY OF REPORT: APPROVED FOR PUBLIC RELEASE; DISTRIBUTION UNLIMITED.
2b. DECLASSIFICATION/DOWNGRADING SCHEDULE	
4. PERFORMING ORGANIZATION REPORT NUMBER(S)	5. MONITORING ORGANIZATION REPORT NUMBER(S): AFIT/CI/CIA-88-222
6a. NAME OF PERFORMING ORGANIZATION: AFIT STUDENT AT ROCHESTER INSTITUTE OF TECH	6b. OFFICE SYMBOL (If applicable):
6c. ADDRESS (City, State, and ZIP Code)	7a. NAME OF MONITORING ORGANIZATION: AFIT/CIA
	7b. ADDRESS (City, State, and ZIP Code): Wright-Patterson AFB OH 45433-6583
8a. NAME OF FUNDING/SPONSORING ORGANIZATION	8b. OFFICE SYMBOL (If applicable):
8c. ADDRESS (City, State, and ZIP Code)	9. PROCUREMENT INSTRUMENT IDENTIFICATION NUMBER
	10. SOURCE OF FUNDING NUMBERS: PROGRAM ELEMENT NO. / PROJECT NO. / TASK NO / WORK UNIT ACCESSION NO.

11. TITLE (Include Security Classification) (UNCLASSIFIED)
LAND COVER CLASSIFICATION OF LANDSAT THEMATIC MAPPER IMAGES USING PSEUDO INVARIATION FEATURE NORMALIZATION APPLIED TO CHANGE DETECTION

12. PERSONAL AUTHOR(S)
CAPTAIN TIM HAWES

13a. TYPE OF REPORT	13b. TIME COVERED	14. DATE OF REPORT (Year, Month, Day)	15. PAGE COUNT
THESIS/~~DISSERTATION~~	FROM _____ TO _____	1988	114

16. SUPPLEMENTARY NOTATION
APPROVED FOR PUBLIC RELEASE IAW AFR 190-1
ERNEST A. HAYGOOD, 1st Lt, USAF
Executive Officer, Civilian Institution Programs

17. COSATI CODES		
FIELD	GROUP	SUB-GROUP

18. SUBJECT TERMS (Continue on reverse if necessary and identify by block number)

19. ABSTRACT (Continue on reverse if necessary and identify by block number)

DTIC ELECTE FEB 01 1990 S D

90 01 31 001

20. DISTRIBUTION/AVAILABILITY OF ABSTRACT: ☒ UNCLASSIFIED/UNLIMITED ☐ SAME AS RPT. ☐ DTIC USERS	21. ABSTRACT SECURITY CLASSIFICATION: UNCLASSIFIED
22a. NAME OF RESPONSIBLE INDIVIDUAL: ERNEST A. HAYGOOD, 1st Lt, USAF	22b. TELEPHONE (Include Area Code): (513) 255-2259 / 22c. OFFICE SYMBOL: AFIT/CI

DD Form 1473, JUN 86 Previous editions are obsolete. SECURITY CLASSIFICATION OF THIS PAGE

AFIT/CI "OVERPRINT"

LAND COVER CLASSIFICATION OF LANDSAT THEMATIC MAPPER IMAGES USING PSEUDO INVARIANT FEATURE NORMALIZATION APPLIED TO CHANGE DETECTION

by

Captain Tim Hawes

B. S. United States Air Force Academy

(1981)

A thesis submitted in partial fulfillment
of the requirements for the degree of Master of Science
in the Center for Imaging Science in the
College of Graphic Arts and Photography of the
Rochester Institute of Technology

September, 1987

Signature of the Author Tim Hawes

Accepted by R. Maxwell 2-12-88
Coordinator, M.S. Degree Program

CENTER FOR IMAGING SCIENCE
COLLEGE OF GRAPHIC ARTS AND PHOTOGRAPHY
ROCHESTER INSTITUTE OF TECHNOLOGY
ROCHESTER, NEW YORK

CERTIFICATE OF APPROVAL

M.S. DEGREE THESIS

The M.S. Degree Thesis of Captain Tim Hawes
has been examined and approved by the thesis committee
as satisfactory for the thesis requirement for the
Master of Science Degree.

Dr. John R. Schott, Thesis Advisor

Dr. Edward M. Granger

Lt Col. Charles P. Datema

2/12/96
Date

THESIS RELEASE PERMISSION FORM

ROCHESTER INSTITUTE OF TECHNOLOGY
COLLEGE OF GRAPHIC ARTS AND PHOTOGRAPHY

I, Tim Hawes, prefer to be contacted each time a request for reproduction is made. I can be reached at the following address:

Tim Hawes
PO Box 624
West Chester, PA 19380

Signature Tim Hawes

Date 18 JAN '88

LAND COVER CLASSIFICATION OF LANDSAT THEMATIC MAPPER IMAGES USING PSEUDO INVARIANT FEATURE NORMALIZATION APPLIED TO CHANGE DETECTION

by
Tim Hawes

Submitted to
the Center of Imaging Science
in partial fulfillment of the requirements for
the Master of Science Degree at
the Rochester Institute of Technology

ABSTRACT

A radiometric normalization technique for compensating illumination and atmospheric differences between multi-temporal images should allow classification of the images with a single classification algorithm. This allows a simpler approach to land cover change detection. Land cover classification of Landsat Thematic Mapper Imagery with and without Pseudo Invariant Feature Normalization was performed to demonstrate the effect on classification and change detection accuracy. A post-classification change detection method using two separate classification algorithms, one for each date, was performed as a baseline comparison. Land cover classification using one classification algorithm was attempted with and without gain and offset correction to serve as another comparison. Accuracy verification was performed on the classification results by comparing random samples against ground truth.

Theses

Acknowledgments

I wish to thank all the members of the Digital Imaging and Remote Sensing Lab for their support and help in the various aspects of this project. In particular, I want to acknowledge:

Carl Salvaggio whose timely advise, help and encouragement during the entire course of my work was much appreciated,

Rolando Raqueno who helped in creating and preparing the programming required for this thesis,

Steven Schultz who aided in the problem solving required in any new programming effort,

The late Dr. Ronald Francis whose guidance transcended his time with us,

And lastly, Dr. John R. Schott whose advice and insight helped keep me on track as the thesis evolved and without whom this thesis would not have occurred.

DEDICATION

This thesis is dedicated to

my wife **Kathy**

who was forced to endure the

trials and tribulations of the thesis process

and continued to support me through it all.

(I am not sure which one of us is more relieved that

the struggle is over.)

TABLE OF CONTENTS

List of Tables

List of Figures

1.0. INTRODUCTION

"Remote Sensing is the acquisition of information about an object without physical contact."[1] This simple definition is expanded by Lillesand (1978) who calls it in combination a science and an art, "...of obtaining information about an object, area, or phenomena through the analysis of data acquired by a device that is not in contact..."[2] Originally termed by naval geographers, remote sensing has grown to encompass any gathering and processing of information about the earth's environment through the use of noncontacting sensors.[1] This thesis deals in particular with the interpretation of satellite collected electromagnetic radiation and specifically with data gathered by the NASA Landsat satellites in the wavelength range from 0.45μm to 2.5μm.

Land cover change detection, simply stated, is the ability to compare two images of the same ground scene, taken at two different times, and determine what land cover has changed in the scene from one time to the next. The ability to accomplish this task on remotely sensed data is made more difficult by variable factors which include; illumination, atmospheric, sensor, and scene variables that can change between image acquisitions and affect the resulting image.[3] Changes in these variables affect the radiance values recorded at the sensor and subsequently the image viewed on the ground. Accurate change detection must account for these variables, but to this point in time there has been limited success in the remote sensing community in quantifying and accurately correcting for these variations to

allow for accurate, repeatable change detection. The ability to detect land cover changes in remotely sensed images of the same target area taken at different times has become one of the many challenges of the remote sensing community.

1.1. Objective.

As introduced above, the problems of differing illumination angles along with variable atmospheric effects have made comparison of multitemporal images, even from similar sensors, difficult if not impossible. The application of the Pseudo-Invariant Feature (PIF) normalization technique now allows compensation for illumination angle and atmosphere differences between images taken at two different times providing they contain some form of constant or pseudo-invariant feature such as an urban area. The PIF normalization allows use of one land cover classifier on both images thereby decreasing the classification time and complexity. This method is compared here for accuracy with a standard method of using a different classifier on each of the two different days, a technique which has been demonstrated in many other studies. In addition to these two methods, a classification of both scenes with one classification algorithm will be performed without using any corrections for atmosphere, sun angle, view angle or sensor. This will also provide a comparison for the PIF normalization technique. Since both test images were taken by Landsat sensors at roughly the same time of day, the effects of differing view angle and sun angle

should be small. This leaves atmospheric and sensor differences between the images. To account for some of the sensor differences, this thesis will also make a comparison classification after correcting for the gain and offset differences between the two Landsat sensors.

Some background information will be reviewed followed by a description of the work performed as a part of this thesis. Then the results obtained by applying the four different change detection variations will be discussed along with areas of possible future investigations.

1.2. Background.

In the following review of current literature, change detection techniques described in the literature will be examined to explore what techniques are applicable to this specific situation. The different spectral bands available from the Landsat Thematic Mapper will be examined to determine which bands are optimal for use in this case of change detection. A review of some selected land cover classification techniques is also included here as the accuracy of the image classification directly affects the accuracy of the change detection. The importance of atmospheric compensation or normalization is discussed with reference to work published in the current literature and its possible application to this proposal. Finally, some suggested methods for determining land cover map accuracy are examined with an evaluation

of possible methods for use in this research.

1.2.1. Change Detection.

In this section, a review is made of some previous change detection examples. To start, three basic approaches are discussed, progressing from a subtraction of corresponding image pixels in each scene, to using a linear relationship derived between the pixels in each scene for normalization, to a subtraction of images after they have been classified according to land cover. Further, an example of an image difference technique combined with gray level thresholding is examined, followed by a look at four variations in land cover change detection that were actually applied to imagery by Weismiller et al (1977). [4]

Jensen et al (1983) discuss three approaches to change detection; 1) image differencing, 2) normalization, and 3) post-classification.[1] A simple form of change detection is a subtraction process where two images of the same scene are registered to each other and then subtracted, pixel by pixel forming a difference image that contains the areas of change between the two images. To explain it another way, image differencing, is a method where a difference value, D_{ij}, is calculated for each pixel (with image coordinates of i-line, j-column) in the image. D_{ij} is calculated by subtracting the radiance

value, X_{ij}, received at the sensor for that pixel at a time t_1 from the X_{ij} value at a time t_2 such that:

$$D_{ij} = X_{ij}\ (t_1) - X_{ij}\ (t_2) \qquad (1)$$

A condition of zero change in radiance from time 1 to time 2 results in a zero value for D_{ij} and is interpreted to mean no change of land cover. The value of one half the dynamic range (for example 128 of 256 total digital counts for 8 bit data) is usually added to the result to eliminate negative numbers. Two problem areas with this method occur: 1) misregistration between the images can cause changes to be detected that did not actually occur; and 2) choosing an appropriate threshold to discern between unimportant and important radiance changes that constitute an actual change in land cover of the scene can be difficult. This method does not adjust for the variables of changing view angle, sun angle, and atmosphere. This thesis will investigate some methods for compensating these variables.

A second approach, mentioned by Jensen et al (1983), to change detection uses a variation of the image differencing mentioned above.[1] In this variation, a least squares transform (image regression performed pixel by pixel) is used to reduce the effects of different atmospheres and/or sun angles between images. The linear model is:[1]

$$e_{ij} = X_{ij}(t_2) - (a + b\,X_{ij})(t_1) \qquad (2)$$

Where X_{ij} is the radiance value at the sensor for time 1 or time 2. "a" is the intercept, "b" is the slope factor that adjusts for differences in variance of radiance values on different dates, and e_{ij} is the residual of the regression. These values are obtained by regression of the individual pixels for time 1 with the corresponding pixels for time 2. The regression accounts for differences in the mean and variance values for different dates. However, this method degrades some of the land use change detection accuracy while reducing the atmospheric and sun angle effects because it assumes that differences between the corresponding pixels are linearly related which may not be true in each image.[1]

A third approach suggested by Jensen et al (1983) involves a post-classification technique which applies a classification algorithm to each of the two dates separately before performing the difference calculation.[1] The classified images can readily be compared and offer the advantage of minimizing the problems associated with differing illumination and atmospheric variables by tailoring the classification algorithm to the specific image. A disadvantage with using this method arises from the high accuracy required from the classification algorithm since any misclassified

pixel may be detected as a changed pixel between the images. Misclassifications can unfortunately occur in either or both of the images, further compounding this problem. Actual land cover changes are detected after the subtraction where they appear as a non-zero value for a given pixel location. This method was used as a standard of comparison in the research.

Schowengerdt (1983) gives a different example of image differencing based on subtracting an image at time one from an image at time two.[5] His method uses gray level thresholding interactively applied by the operator on the difference image to achieve a display of gray level changes that exceed a certain magnitude of change between images. This method, while certainly showing potential, will not be evaluated in this work as it does not involve classifying the image until after the image subtraction process.

Weismiller et al (1977) analyzed four variations of change detection in trying to detect areas of land cover change.[4] The first, a post-classification method, utilized a separate classification algorithm on each individual date before comparison of the two dates. This method was then used by the authors as the basis of comparison for the remaining variations. Likewise in this thesis, a post-classification method was used as a standard for comparison. A second variation, "delta date" change detection, created a subtraction image of the two dates and then applied a classification routine to

the resulting image. The authors dismissed this method for future use as being to simple to account for all the factors involved in the natural scene.[4] A third variation used by Weismiller et al (1977) was based on combining the data from the images of two dates into a single eight channel data set. By choosing the dates to be approximately one year apart, the data could be classified using "standard pattern recognition techniques". Areas that contained no change between the images would have different statistics than those areas that contained change from one time to the next. This method has the disadvantage of requiring specific timing of the images to be compared and limits the applicability of the technique. The authors did feel that this method deserved further research, but was not explored in this thesis. Lastly, a fourth variation, "the layered, spectral/temporal approach," applied a maximum-likelihood classifier with multi-stage decision logic to target areas on each of the different dates.[4] The results did not show a significant increase in accuracy over the post-classification change detection method and was not explored any further by this thesis.

In this section, different approaches to change detection and their limitations were examined. The post-classification method was chosen for use in this research over the other three methods. Image subtraction of the classified images was also choosen to detect land cover changes. The post-classification technique combined with image subtraction method of land cover change detection was

selected as the standard of comparison for this research with a variation of the normalization method used to demonstrate the advantages available from normalizing the two images.

1.2.2. Band Selection.

With the increasing capability of modern sensors to collect large amounts of data over many wavelengths, the emphasis of change detection has shifted from gathering and using all the information available to a situation where only the amount of data necessary to provide accurate information is used. The goal of spectral band selection is to obtain the most information about the scene with the least amount of data processed.

In the case of the Landsat Thematic Mapper (TM), the spectral response for any particular object is composed of seven spectral bands. Although these bands do not spectrally overlap, they contain information that overlaps to some extent. This information overlap can cause confusion in the identification process and increases the complexity and time required to make the correct classification. The key is to find the selection of bands that have enough information to allow assignment of a spectral return to a particular land cover class, but not so much information that the process is slowed unnecessarily or confused. The following section reviews some of the relevant literature that discusses selecting spectral bands to limit

information redundancy while retaining identification accuracy.

Toll (1985) in a comparison of overall classification accuracy for Landsat TM imagery of Laurel, MD found, "...that middle infrared bands provide the best spectral information to discriminate among classes."[6] He also found that the visible TM bands 2 and 3, "...provided the next best discriminator." This was using a per pixel Gaussian maximum likelihood classifier.

Toll (1985) continued his analysis with multiband combinations and found that the best spectral discriminations occurred when spectral bands from the visible, near infrared, and middle infrared were included. He found that band 2 information is highly correlated with bands 1 and 3. Also, band 7 is highly correlated with band 5. Toll (1985) concludes that bands 2 and 7 are the first bands that should be excluded, if required. He found too, that excluding them did not significantly affect the classification accuracies.

Gervin et al (1958) developed similar conclusions. They maintained that 3 or 4 band combinations with at least one band in the visible, near infrared, and middle infrared regions provided, "... nearly as good a discrimination of land cover as all seven bands."[7]

In a separate study concerning the selection of 3 band combinations, Sheffield (1985) attempted to determine what 3 bands

provided the most information about a 7 band TM scene.[8] He found that the combination selected is somewhat dependent on the original scene, but, "The band combination 1, 4, 5 ... is usually, but not always, the selected triplet. In cases where it does not rank first, it usually ranked second." He went on to state that the selected bands usually came, one each, from the visible, near infrared, and short wave infrared (middle infrared). Also, he found the triplets that ranked highly, always included band 5 or band 7. This further confirms the importance of the information contained in the middle infrared region.

An English study, by Dean et al (1986) further supports the findings of the previous works.[9] Using a maximum likelihood classifier, based on field observed training areas, they found the most reliable results to be gained from using TM bands 3, 4, 5, and 7.

The use of thermal infrared data was not largely encouraged by the previous studies, but one author went so far as to discourage its use. Price (1981), in an article on the contribution of thermal data in Landsat multispectral classification, contends that although the thermal information may be independent of the reflected radiation bands, "... the numerous physical processes governing thermal radiation lead to a dependence on surface slope, altitude, and surface energy-balance effects such as ground heat flux, atmospheric heating, and surface evaporation. These effects do not influence the spectral behavior in the reflected channels."[10] He concludes that these

factors make analysis difficult and confuse the classification problem.

The conclusion drawn from this selection of literature is that four spectral bands of information containing at least one band from the visible, the near infrared and the middle infrared should provide enough information to allow accurate classification of land cover classes. For these reasons, the proposed research intended to use TM bands 1, 3, 4, and 5 for classification algorithms, but to increase classification accuracy as much as possible, band 7 was later added to the classification process.

1.2.3. Classification Techniques.

A large variety of classification techniques have surfaced in the literature. These techniques fall into a general category of either unsupervised classification, where the computer determines the land use classes based on statistical groupings, or supervised classification where the operator provides class identification data in the form of previously identified land use training sites. Within these categories lie maximum likelihood classifiers, minimum distance classifiers, use of spatial and contextual information, factor and vector analysis, and techniques using combinations of the previous techniques, such as ECHO[11]. Since many of the change detection techniques are sensitive to the accuracy and consistency of the classifier used, the choice of classifier has a large effect on the final change detection. Following is a review of some of the techniques currently being used or suggested for land use classification:

Jackson et al (1980), faced with an operational land use classification task, looked at three options for classifying Landsat images according to land use, "...(1) manually, (2) with computers plus manual assistance in the form of supervised classification algorithms, or (3) fully automatedly using unsupervised clustering techniques."[12] They chose to use manually defined training sites which then were identified using their principal component axes. The

Landsat data was then classified using the principal component training set data with an assumed Gaussian maximum likelihood. With this classification of single date Landsat MSS data, they achieved accuracies of 70% to 80% correctly identified land use classes. In an attempt to reduce this error, they used a second date of Landsat MSS imagery which had been classified using the same technique to compare with the first classification. The results of this procedure, claimed as successful, were given only as a general reduction in acreage classified as developed land. The authors admitted that the above methods are too simplified for their tasks and are being refined for future use. In view of the success achieved by the use of training sites and supervised classification, the supervised technique was selected as the method of classification for this research.

Tom and Miller (1984) compared the Bayesian maximum likelihood (BML) to the linear discriminant analysis (LDA) algorithm for application to automated land use mapping.[13] The authors pointed out that, "... the maximum likelihood decision rules require a large number of multiplication and logical comparisons for each decision, particularly when many MSS channels and mapping classes are used." They hypothesized that a supervised approach, the LDA, using a priori data in the form of known training sites would offer speed gains (and hence lower computer costs) without lowering classification accuracy. This hypothesis was tested on a subscene of Denver, Co.

with the following results reported:

- Linear discriminant analysis was a more accurate classifier than the Bayesian maximum likelihood
- Bayesian maximum likelihood machine time was greater than that for linear discriminant analysis
- Bayesian maximum likelihood machine cost per correctly classified pixel was much greater than that for linear discriminant analysis; and
- Linear discriminant analysis machine time was much less sensitive to the number of mapping variables and zero mapping class variance than Bayesian maximum likelihood.

The linear discriminant function as it is applied in the DIRS Gould image processing system functions as a "minimum distance to the mean" algorithm. The training set data provided by the supervised portion of the computer routine provides a set of class means which are represented as vectors. The class variances are assumed to be equal and the pixel in question is assigned to the class with the closest mean measured by Euclidean distance. In view of these speed advantages, linear discriminant analysis was selected as the classifier for this research.

Hsu (1978) investigated the use of texture-tone for classification in land use mapping.[14] In work on black and white imagery, he was searching for ways to separate classes based on the spatial distribution of the image data. He defined texture to mean,"... the spatial distributions of tones of the pixels." and chose to measure it statistically. His technique utilizes a 3 x 3 window forming a

"texture feature extractor," as well as a 5 x 5 window. These two windows allowed calculation of different statistical data such as the first four central moments, the absolute deviation from the mean, mean brightness of the center point relative to the background, the square of the previous value, the contrast of the center point from its neighbors, and more, up to 23 variables. Hsu also chose to use a linear discriminant analysis with a Mahalanobis classifier instead of a Bayesian maximum likelihood approach. This model was applied to black and white imagery yielding 85% to 90% correct classification for seven test areas in New York State. The author is quick to point out that this same technique could be combined with spectral information for even greater accuracies, but it was not demonstrated. This method is, unfortunately, computationally intensive and takes large amounts of computer time. For that reason, this approach was not appropriate for this research. A standard deviation image was made using a 3x3 window for this thesis, but it did not add any accuracy to the linear discriminant analysis performed and so was not used.

In an attempt to utilize some of the spatial information in an image along with the spectral information, Kettig and Landgrebe (1976) developed a concept to, "...exploit a particular type of dependence between adjacent states of nature that is characteristic of the data."[11] They call this technique, "Extraction and Classification of Homogeneous Objects" or ECHO. Their technique is

based on the assumption, "... that each object is represented by an array of pixels. This suggests a statistical dependence between consecutive states of nature." They propose a partitioning of the image into arrays of statistically similar pixels before applying a classification algorithm. This results in an image that has been classified according to the spectral information of the neighbors of a pixel as well as the pixel's own spectral content. This technique proved to have a consistently lower error rate when applied to each of the authors' four test data sets over the conventional maximum likelihood approach. Another advantage is the reduction in number of classifications necessary, thus decreasing computer time required for the classification step. This technique seems useful for change detection classification, but needs further investigation and was not used as part of this research.

A method for classifying images using spatial context information was proposed by Swain et al (1981).[15] Their concept is based on the assumption that certain land cover classes are more likely to occur in the company of other land cover classes. As they state it, "One does not expect to find wheat growing in the midst of a housing subdivision, for example. A close-grown lush vegetative cover in such a location is more likely the turf of a lawn." The key is to be able to accurately define the relationships between the classes, the "context distribution". The authors tested their theory on some simulated image data and on some real image data. The results were

encouraging, but the authors concede further work needs to be accomplished and that the computational requirements were currently large. For these reasons, the method was not considered for use in this research.

Huete (1986) explored a novel technique of separating soil- plant spectral mixes by factor analysis.[16] In his laboratory demonstration, the author was able to decompose a data set of spectral mixtures, "... into the sum of unique reflecting components weighted by their corresponding amounts." In other words, he was able to separate a spectral return into the components that caused specific portions of the return. In this case, variable amounts of plant densities and soil mixtures were decomposed into dry soil, wet soil and vegetation components. This breakdown was accomplished using matrices to categorize the spectral mixture into identifiable components. This technique has potential for allowing comparison of results over different locations and allowing search for components of interest using spectral signatures. Although this research has a potential for future use, it is not developed far enough for use in this thesis.

Carr et al (1983), used supervised classification of Landsat images to classify Arizona copper mining activity.[17] They were attempting to extend the spectral signatures developed for one mine to be used in the classification of other mine sites in the Arizona area. This required both a temporal and a spatial extension by taking

many samples at one mine site to develop a range of signature values for application at the other mine sites without retraining the classifier. The authors successfully accomplished this after normalizing each scene with a solar correction factor and by excluding the band 4 (MSS green wavelengths) data due to atmospheric effects. This study further supports the use of supervised classification over unsupervised in areas where training sites are known and explored the concept of building class mean values (or signatures) to apply to normalized images in different locations. A supervised post-classification method was used in this thesis as well as a variation of the image normalization technique to demonstrate the advantages found with these methods.

Sanchez and Taheri (1985) developed a preprocessing technique for unsupervised clustering of spatially and spectrally similar image pixels.[18] They simply included the spatial coordinates of the pixel in a vector with the spectral information about the pixel. The pixels with similar spectral and location vectors could then be grouped together to form "blobs". These "blobs" were then assigned a mean vector and the classification process was accomplished using the mean vectors. This clustering should shorten the computer time required to classify the image, but the creation of the "blobs" would add some preprocessing time. Unfortunately, the authors did not demonstrate the technique on real world data or have any accuracy figures. For these reasons this technique was not used in the thesis

research.

The accuracy and complexity of unsupervised classifiers currently dictate the selection of a supervised technique for this research. The accurate results of linear discriminant analysis make this technique a good choice to use in combination with a supervised classification.

1.2.4. Accuracy Verification.

In order for the change detection map to be useful, it is necessary to asess the accuracy of the changes identified, "...with the collection of some in situ data..." or ground truth.[1] Since the entire scene cannot be realistically verified due to its large size, the concept of sampling is used to limit the amount of ground work required. From an efficiency and cost standpoint, the number of samples must be large enough to statistically verify the accuracy of the changes detected and still minimize the ground verification required. These constraints are discussed in the following articles.

Van Genderen and Lock (1977) discuss a sampling strategy for minimizing the number of samples required (per land cover class to be tested) and yet still maintain statistical validity.[19] The number of samples required is based on a binomial expansion given the interpretation accuracy required. The results are shown in an error matrix format to allow spotting of trends (see table I-1). They claim

that their method is applicable to orbital data and indeed any image interpretation based land-use classification. This approach seems suitable for a "first look" at the classifier results in the proposed research.

TABLE I-1

ERROR MATRIX EXAMPLE

		LAND USE ON GROUND			
		A	B	C	SUM
LAND USE FROM IMAGERY	A	21	3	6	30
	B	2	28	0	30
	C	1	2	27	30
	SUM	24	33	33	90

Rosenfield et al (1982) have also developed a minimum sample size method to test the accuracy of a classification algorithm.[20] Their system is developed from the cumulative binomial distribution and is tied to a specified level of confidence (ie. 95%). They have developed a computer program that will perform a sampling technique with an additional random sample of points for under-represented categories. The computer will also accept field identified classifications for an automatic comparison with the interpreted classifications. A printed copy of this program was obtained from the United States Geological Survey for evaluation of its usefulness to this project. Unfortunately, only the concept of automated random test point selection was adaptable for the current work as the bulk of the computer software required to run the program was not available at the DIRS laboratory.

Todd et al (1980) demonstrated a random sampling strategy in their assessment of wildland mapping accuracy using the Landsat MSS.[36] They used the following formulae to calculate the minimum number of samples required for each class:

$$n_i = \frac{N_i \; p_i \; q_i}{N_i \left(E^2 / t^2 \right) + p_i \, q_i} \qquad (3)$$

where:

n_i = the sample size of class i

N_i = the total size of class i

p_i = the estimated accuracy of class i (provided by ground truth estimation)

$q_i = 1 - p_i$

E = the allowable error (ie. 10%)

The percentage correct for each class was calculated using:

$$C_j = \frac{\Sigma b_{jk}}{\Sigma a_{jk}} \qquad (4)$$

where:

C_j = the percent correct for class i

b_{jk} = *number of observations correctly classified into* class j in sample k

a_{jk} = number of observations of class j in sample k actually found in ground truth

This allowed calculation of the standard error of the estimate C_j:

$$E(C_j) = \frac{(\sqrt{1-h})}{[(\sqrt{n})(\Sigma a_{jk})/n]} \cdot \sqrt{\frac{/\ \Sigma[b_{jk} - (C_j/100)\ (a_{jk})]^2}{(n-1)}} \qquad (5)$$

where:

in addition to the variables defined above,

$$h = n_j/N_j$$

From all the above, the statistical confidence interval can be calculated for each class:

$$I_j = C_j \pm (t) [E (C_j)] \quad (6)$$

where:

I_j = Confidence interval for class j

t = Appropriate critical value of the Student's t distribution

This article demonstrated a statistical approach to assessing the accuracy of any classifier independent of the approach used to obtain the classifier. A variation of this approach was examined for use in evaluating the classifiers created during this research, but the following method proved simpler.

Hord and Brooner (1976) approach land-use classification accuracy from an equally statistical perspective, basing their procedure on the binomial distribution applied to a random sample of the classes. They approximated this distribution using:[22]

$$P\left(-b < \left(\frac{X - \mu}{\sigma - N}\right) < b \right) = 1 - \partial$$

where:

N = the number of samples in the class

x = the sample mean

ø = the sample standard deviation and

$\partial = 0.05$ (for a 95% confidence level)

$b = 1.960$ (from the normal distribution)

This equation is then solved for the upper and lower limits of μ. This method was used to demonstrate the application of confidence intervals in this thesis.

The authors suggest that the minimum sample required to use these assumptions is 50 samples. Hay (1979) insists that at least 50 samples per class are necessary for accurate sampling in his article on sampling designs for use in testing land-use map accuracy.[23] He uses the Binomial distribution assumption, but only with a minimum of 50 random samples per class to define confidence intervals for the sampling results. The use of the Binomial distribution assumption is widely supported and is used in this thesis. The accepted lower limit of at least 50 samples per class was used as a parameter for this thesis to allow analysis of a 95% confidence interval.

There are even accuracy verifications possible without field samples or so Adeniyi (1985) claims in his study of multispectral Landsat data in Nigeria.[24] He was forced to develop a non-sampling approach due to a lack of temporal ground truth data. Instead he used:

- coefficient of variation for within class variability
- two tail test concerning the means of the spectral

signature of the training areas used in defining each class; and

- confusion and divergence matrices.

This method does not appear to be applicable to this research and was not used.

For this research, the method of Hord and Brooner (1976) will be used as it simplifies the accuracy verification process giving the statistical results necessary at a minimum number of sample points to be checked.

1.2.5. Normalization.

Various methods have been proposed for correcting temporally varying images for atmospheric effects and illumination differences. Gerson et al (1983) examined five techniques for temporal image normalization.[25] These techniques included 1) an analysis of image and scene content and/or 2) analysis of atmospheric and ephemeris data to develop algorithms that help remove effects of sun angle, look angle and atmosphere from images of the same scene. After testing, the authors claimed to have obtained good results for black and white aerial photos, but poorer results for color imagery. They concluded that temporal normalization is affected by scene content, the imaging process, and the "perturbation process", effects causing the need for normalization. With none of the five methods tested did they find a

general normalization algorithm.

Volchok and Schott (1985) demonstrated a technique for multitemporal image normalization based on the Pseudo-Invariant reflectance characteristics of concretes and asphalts.[3] This technique was developed from earlier research done by Piech et al (1977).[26] They showed that, "the reflectance of concretes and asphalts, integrated over the visible wavelengths, varied on the order of one reflectance unit." for temporal change.[3] This work was combined with the efforts of Egbert and Ulaby (1972) who demonstrated the constancy of reflectivity for asphalt samples for solar altitudes of greater than 30^{o} and the work of Grogan (1983) with concrete reflectances over varying illumination angles.[27,28] Out of these varied efforts, Volchok and Schott (1985) concluded that, "... asphalt and concrete features can be considered as having approximately invariant reflectance characteristics with respect to time and illumination angle." They labeled them Pseudo-Invariant Features (PIF's). One assumption included here is that the proven visible radiation invariant characteristics extend into the reflected infrared portion of the spectrum. Extension of this assumption into the thermal infrared region is unwarranted.

The invariant reflectance of the concrete and asphalt targets allowed the authors to support the theory, "... that a linear relationship exists between the spectral brightness distribution of

the PIF's in one scene and the spectral brightness distribution of the PIF's in a second."[48] They then used this linear relationship to develop the transform function for scene normalization.

The concept of constant reflectance targets was also used by Lo, Scarpace and Lillesand (1986) in their study of Landsat Multispectral Scanner (MSS) data.[29] In this case, sample pixels of constant reflectance with time, "...(such as airport runways)..." were used to calculate an additive atmospheric haze/path radiance component that was calculated for (and subtracted from) each band of the imagery. This was a less sophisticated atmospheric correction using constant reflectance targets. Unfortunately, the authors did not give any results indicating the effect of the normalization.

For this research, the PIF technique was used to provide image normalization for images of dates over two years apart. The ability of this normalization method to allow use of one classification algorithm for classifying both images will be compared to three other classification methods for the ability to compensate for differences in atmosphere, view angle, sun angle, and sensor. This will be measured by the difference in land use classifier accuracy.

1.2.6. Gain and Offset Adjustment.

As each Landsat sensor stabilizes on orbit, the detectors are

claibrated to an onboard source. Markham and Barker (1986) note that, "...calibration constants... have been modified after launch either because of inconsistency within or between bands, changes in the inherent dynamic range of the sensors or desires to make QCAL values form sensor match those from another."[30] To account for some of the sensor differences, the thesis research will also make a comparison classification after correcting for the gain and offset differences between the two Landsat sensors. This correction will be based on the following:

The output of each sensor gain can be characterized by the following equations:

$$RAD_1 = Lmin_1 + \frac{(Lmax_1 - Lmin_1)}{(Qcalmax)}(DC_1) \quad (7)$$

$$RAD_2 = Lmin_2 + \frac{(Lmax_2 - Lmin_2)}{(Qcalmax)}(DC_2) \quad (8)$$

Where:

DC = the Digital Count output for the sensor (1 or 2)
RAD = the radiance received at the sensor (1 or 2)
Lmin = the spectral radiance at DC = 0
Lmax = the spectral radiance at DC = 255
Qcalmax = the range of rescaled radiance in DC

When looking at a known calibration source where the radiances at the sensor are equal:

$$RAD_1 = RAD_2$$

The two equations can then be set equal to each other so that:

$$Lmin_1 + \frac{(Lmax_1 - Lmin_1)}{255}(DC_1) = Lmin_2 + \frac{(Lmax_2 - Lmin_2)}{255}(DC_2) \quad (9)$$

And finally rewritten so that the Digital Count of one sensor can be corrected for the differing gain and offset of the other:

$$DC_1 = 255 \times \frac{(Lmin_2 - Lmin_1)}{(Lmax_1 - Lmin_1)} + \frac{(Lmax_2 - Lmin_2)}{(Lmax_1 - Lmin_1)}(DC_2) \quad (10)$$

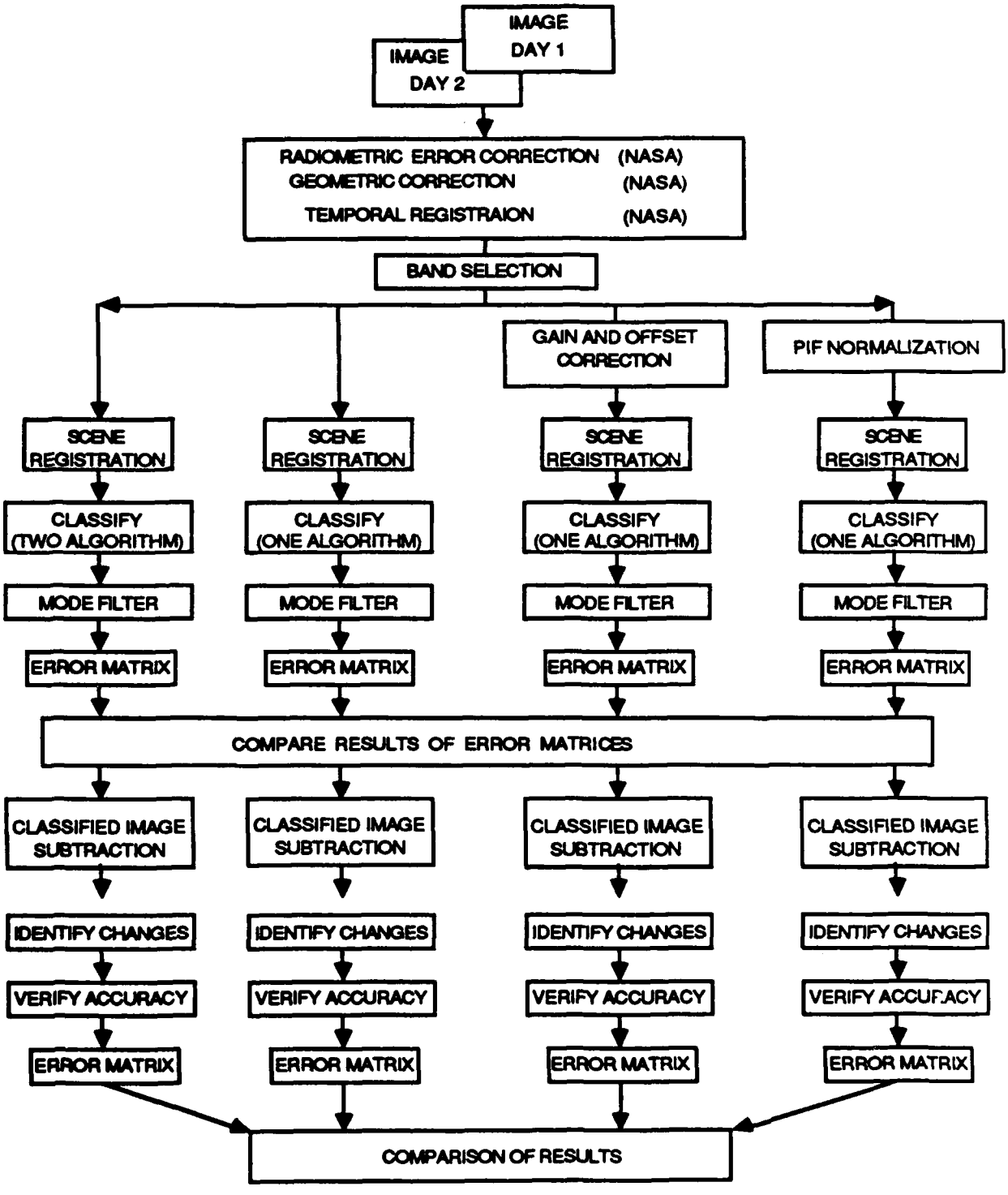
This gain and offset correction will serve as a first order correction for sensor differences and enable a comparison with the PIF normalization technique which should also account for these differences as well as for atmospheric differences.

2.0. METHOD.

The experimental work performed for this thesis included image selection, band selection, gain and offset correction, PIF normalization, scene registration, scene classification, random sampling of classified images, ground truth comparisons, error analysis, classified image subtraction and a final analysis of the results. The same ground scene was used for five different image classifications which were then used for four different change detection tests. The flow of the experiment is illustrated in figure 1. The selected images were processed in the order depicted in figure 1 with each block explained in detail in the following sections.

Fig. 1

RESEARCH FLOW DIAGRAM



2.1. Multitemporal LANDSAT Thematic Mapper Images.

Landsat Thematic Mapper images were selected for this analysis as they contain seven bands of multispectral information spread from the visible through the reflected infrared to the thermal infrared portions of the spectrum. This diversity of spectral information added flexibility to the selection of land cover classification algorithms yet the imagery is widely available. The following table was taken from Freden and Gordon (1983) and shows the spectral range of the particular TM bands:[31]

TABLE II-1

Thematic Mapper Sensor Bands

BANDS	RANGE (μm)
1	O.45-0.52
2	0.52-0.60
3	0.63-0.69
4	0.76-0.90
5	1.55-1.75
6	10.4-12.5
7	2.08-2.35

2.1.1. Subscene Location

Two subscenes of the Rochester area were chosen for ease in ground truth verification. They are 512 x 512 scenes which contain a mixture of urban and rural features that form a broad range of land use classes. These two subscenes have an added advantage that they have been used in a previous demonstration of the PIF technique. The image data is summarized in table II-2.[32]

2.1.2. Sensor Geometric and Radiometric Corrections (NASA)

The TM imagery data is corrected at the ground processing facility for the following conditions according to Freden and Gordon (1983):

Radiometric Error Correction -	1 quantum level over full range
Geometric Error Correction -	0.5 sensor pixel
Temporal Registration Error -	0.3 sensor pixel

TABLE II-2

TM SCENE LOCATION DATA

SCENE ID #	E-30059-15244	Y-50449-15274
SENSOR	TM-4	TM-5
DATE	13 SEP 82	24 MAY 85
APPROXIMATE TIME	10:30 AM	10:30 AM
LANDSAT SCENE COORDINATES	PATH D-17 ROW 30	PATH D-17 ROW 30
SUN ELEV. ANGLE	44 DEG	58 DEG
SUN AZIMUTH ANGLE	141 DEG	127 DEG

The images were originally obtained from the EROS Data Center in Sioux Falls, SD. They are currently contained in the Digital Image and Remote Sensing (DIRS) Laboratory on computer magnetic tape and on 8" floppy disks.

2.1.3. Band Selection

The seven bands available from TM imagery are an advantage in building spectral signatures of land cover classes, but the large amounts of data to be processed to obtain that information is an obstacle to efficient image processing. As is usual in science, a compromise must be made, in this case between having enough data to correctly identify land cover classes, but not so much data that processing becomes cumbersome. As Gervin et al (1958) showed in their paper, it is possible to obtain, "...as good a discrimination of land cover..." with 3 or 4 spectral bands as is possible with all seven TM bands. The key is to choose at least one band in each prominent region of the spectrum (visible, near infrared and middle infrared). With this research in mind, only bands 1, 3, 4 and 5 were initially utilized rather than all seven bands in order to limit the data manipulation required.

In order to increase classification accuracy, band 7 was later added to the classification process. Also band 7 was used in the PIF normalization process in creating the mask for urban features. The additional band did increase the complexity of the data manipulation somewhat.

2.2. Post-Classification Change Detection (Two algorithms).

The concept of comparing two multitemporal images is eased by limiting the comparison to an arbitrary number of preselected classes, in this case only five. This technique has been used as a basis for comparison in other studies and is appropriate to use here.

2.2.1. Registration of the Two Scenes.

The two Rochester scenes were taken at different sensor orbital positions which caused the scenes to not overlay one another exactly. To correct this problem, 18 control points were selected that were visible in both images (see Appendix A). A multiple linear regression was run to determine the transformation equation that most correctly translated the image control points from one image to the other. The following transformation equations resulted:

$$X = 62.25 + 1.01(X_{82}) + 8.90E\text{-}03(Y_{82}) + 4.61E\text{-}06(X_{82})(Y_{82}) - 2.09E\text{-}05(X_{82})^2 - 1.99E\text{-}05(Y_{82})^2$$

$$Y = 41.91 - 7.08E\text{-}03(Y_{82}) + 1.00(X_{82}) - 2.63E\text{-}06\,(X_{82})(Y_{82}) + 2.00E\text{-}05(Y_{82})^2 + 7.37E\text{-}06(X_{82})^2$$

where (X, Y) are the (row, column) coordinates for the registered 1982 scene and (X_{82}, Y_{82}) are the coordinates of the original 1982 scene. The error from the control points (see Appendix A) indicated

that the error introduced by the registration would only be 0.3 pixels in the X direction and 0.4 pixels in the Y direction.

The bilinear interpolation method was used to calculate the proper digital count values to allow one image to be "overlayed " on it's temporal pair. In this case, the 1982 image was resampled to register with the 1985 image using the resampling algorithms available on the DIRS Gould DeAnza Array Processor. A small amount of radiometric error was introduced by bilinear interpolation as it assigns average values created from surrounding pixels to 1982 pixels that do not completely coincide with existing 1985 pixels. This should have had small effect on the classification process and was assumed to be insignificant.

A visual check of the registration process was performed to verify the results. Most of the resampled 1982 image was indeed within one pixel of the 1985 image, but one section on the east side of the image was more than one pixel out of registration. This same resgistration method was used on all the 1982 scenes before the classification step. This registration error had a large impact on the change detection process as any misregistration can be misinterpretated as a change after image subtraction.

2.2.2. Supervised Training Set Construction.

The land cover classes can be determined many ways, but a supervised, training site selection process was chosen where the operator used known training sites to build spectral signatures. These training sites were easily identified by land use maps, photo interpretation, personal observation or a combination of all three. Some experimentation was conducted with an unsupervised cluster analysis program to identify land cover classes. Unfortunately the accuracy of the classified image was not high enough to qualify the unsupervised approach for use in post-classification change detection. Instead, the spectral signatures for five different classes were built using the supervised training site method and the sampling software available in the LIPS software library owned by DIRS. The class mean vectors are built based upon the pooled variances of the samples. The five classes selected are shown in table II-3 along with the class means for the 1985 image.

TABLE II-3

1985 CLASSIFICATION MEANS

	SELECTED LAND USE CLASS MEANS (1985)				
	Band 1	Band 3	Band 4	Band 5	Band 7
URBAN	121.20	62.39	49.50	78.72	51.76
WATER	88.76	29.73	14.74	9.94	3.76
RESIDENTIAL	102.59	44.13	85.98	75.81	32.34
VEGETATION	84.88	31.15	123.43	88.01	26.48
SOIL	144.27	104.59	105.32	170.95	108.89

2.2.3. Confusion/Error Matrix for Classification.

These class means were then used by the computer to classify the sample points choosen for the training sets as an indicator of the classifier accuracy. Each pixel in the training set was assigned to the class whose mean vector was closest to the actual pixel values, as measured by a linear discriminant analysis in the five dimensional class mean space. The results of the classification of these known sample points is displayed as a confusion matrix in table II-4. This confusion matrix was used as part of an interative selection process

in building the training sets. New additions to the training set were evaluated on the amount of incorrect classification they caused in this matrix and then discarded if they did not increase the accuracy of the classifier. The overall performance percentage was only used as a means of comparing the different combinations of training sets. It is not an absolute measure of the accuracy of the classification of the image. The final class means for 1985 are displayed in table II-3 and shown graphically in Figure II-1.

Table II-4

1985 Classifier Confusion/Error Matrix

	Samples	URB	WAT	RES	VEG	SOIL
						(IN %)
URBAN	224	99	1	0	0	0
WATER	70	0	100	0	0	0
RESIDENTIAL	1032	1	0	96	3	0
VEGETATION	348	0	0	1	99	0
SOIL	273	0	0	0	0	100
TOTAL	1947					

OVERALL PERFORMANCE (%): 97.48

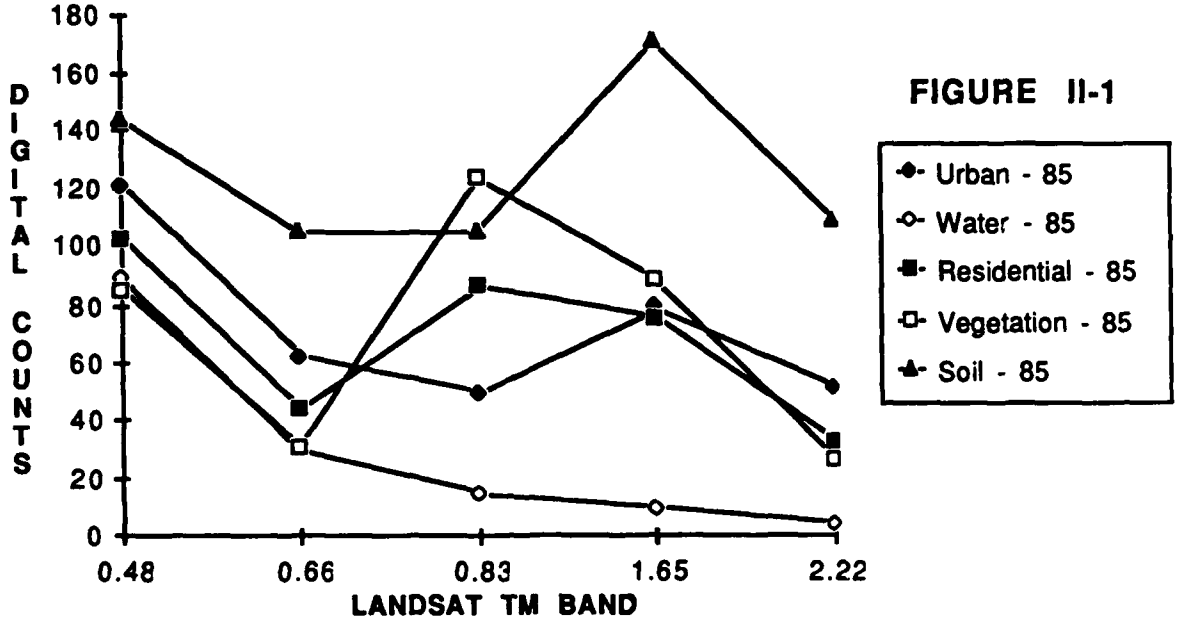


Figure II-1: 1985 Class Mean Vectors

The same training process was performed on the 1982 scene. For some of the classes, an attempt was made to use the same training sites (ground locations) for the 1982 image as had been used for the 1985 image. This worked somewhat successfully for the residential training locations, the water and the urban areas as they had not changed significantly between the two dates. However, the soil and vegetation training sites had undergone significant changes and had to be reselected. The extra atmospheric moisture, or haze, in the 1982 image also compounded this problem by affecting the reflectance values of the known training locations. For example, the classifier had a difficult time distinguishing the airport runways at Rochester

International Airport from a soil land cover. This is paritally due to the high reflectance of the concrete that forms much of the runways, but the haze did not help this situation. The resulting class means for each band in the 1982 scene are found in table II-5 with the information shown graphically in figure II-2.

TABLE II-5

1982 CLASSIFICATION MEANS

SELECTED LAND USE CLASS MEANS (1982)					
	Band 1	Band 3	Band 4	Band 5	Band 7
URBAN	105.26	44.13	41.60	55.61	31.58
WATER	93.19	31.11	24.63	11.20	4.56
RESIDENTIAL	94.37	33.86	56.52	48.97	18.69
VEGETATION	90.64	31.92	76.74	65.13	19.51
SOIL	101.26	49.88	68.25	104.20	53.60

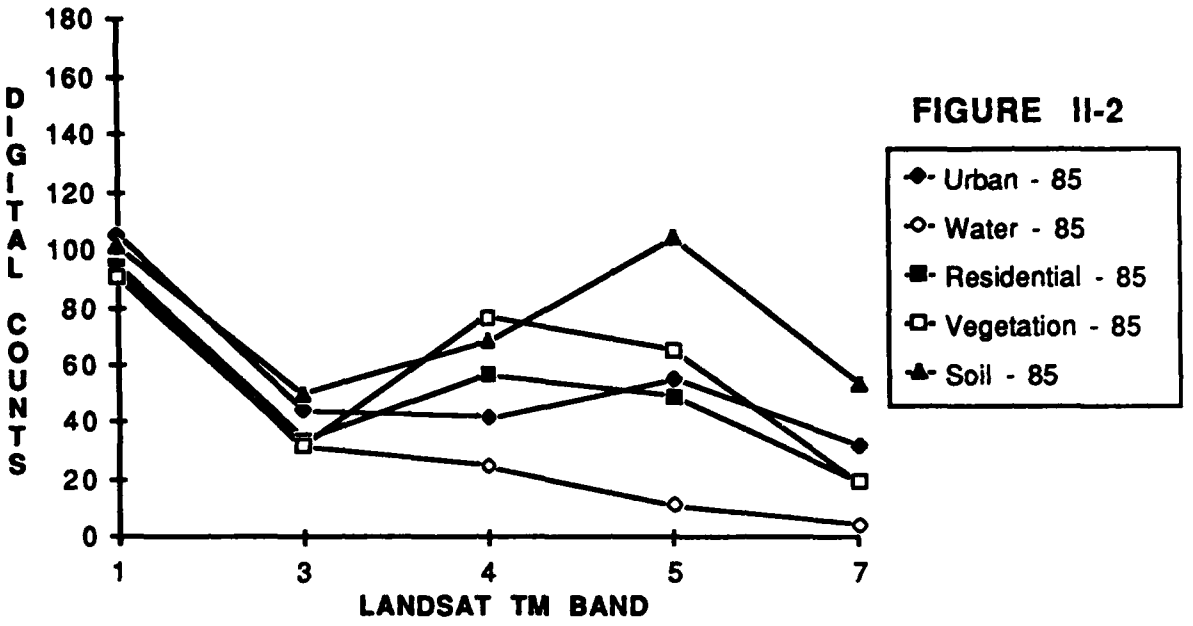


Figure II-2: 1982 Class Mean Vectors

Again, the class means were developed with an iterative use of the confusion matrix to determine what training samples decreased or increased the accuracy of the classifier. The final class means did not perform as well on the 1982 training set as the 1985 class means had performed on their training set, but the training set results were still above 93% correct which indicated that the accuracy was still acceptable. The difference between the two results is partly attributable to the larger training set size required for the 1982 scene which introduced some eroneous pixels, especially around the airport area, but the results were also adversly affected by the haze which lowered the contrast between reflectance values of different training sets. The final results of the confusion matrix are shown in

table II-6.

Table II-6

1982 Classifier Confusion/Error Matrix

	# of Samples	URB	WAT	RES	VEG	SOIL (IN %)
URBAN	348	87	5	0	1	6
WATER	54	0	100	0	0	0
RESIDENTIAL	1559	2	0	94	4	0
VEGETATION	895	0	0	4	93	3
SOIL	81	0	0	0	0	100
TOTAL	3145					

OVERALL PERFORMANCE (%): 93.39

With the completion of the classifiers, the next step was to actually classify the images. The LIPS software routine, "CLAMID" was used to classify the images using a minimum distance to the class mean classifier. This technique assigns the current pixel to the class mean which is closest to it in five dimensional space and assumes the class variances are equal.

The resulting classified images had areas where mixed pixel effects had caused some apparent errors in the classification. Mixed pixel effects occurred along borders of classes where the objects of one class may only partially fill the pixel and hence the sensor assigns that pixel to a wrong class based on the different return. In order to overcome some of these effects, a (3x3) mode filter window was passed over the image to smooth isolated mixed pixel effects. The mode filter was implemented by passing a window over the image that counted all the pixel values contained in the window (9 in this case) and deposited the most frequently occuring pixel value in the center location of the window. The net effect of the filter was to change isolated pixels, which were most likely wrong, to the most frequently occuring value of their neighbors in a 3 pixel by 3 pixel area. This techniqe did introduce some error of its own, and the tradeoff performed was to apply just enough smoothing to reduce mixed pixel effects, but not enough to introduce unacceptable amounts of error. For this application, a 5 x 5 window was also tried, but it caused an unacceptable loss of accuracy in small class areas such as runways and roads. The results in terms of changed pixels are illustrated in figure II-3 for the 1985 image and in figure II-4 for the 1982 image classified with its own class mean vectors.

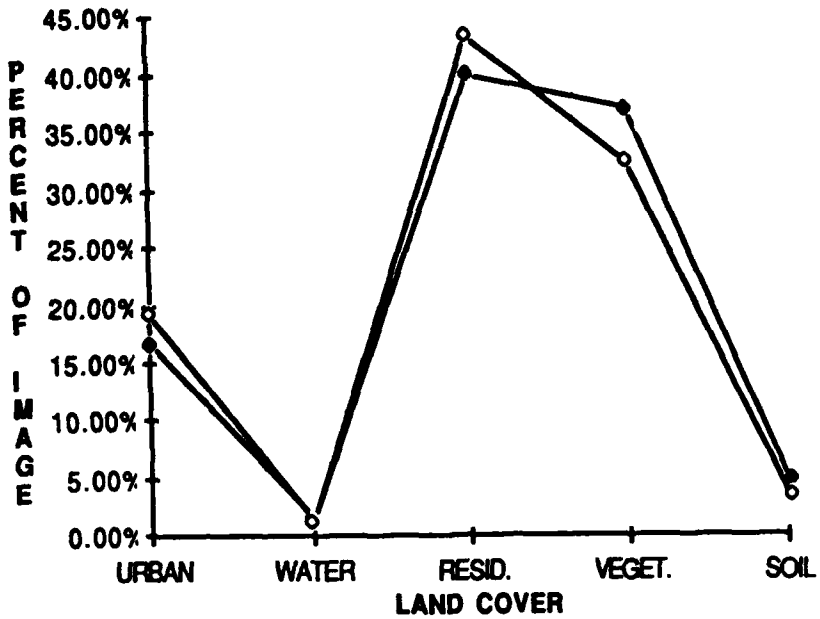


FIGURE II-3

Figure II-3: 1985-Mode Filter Comparison

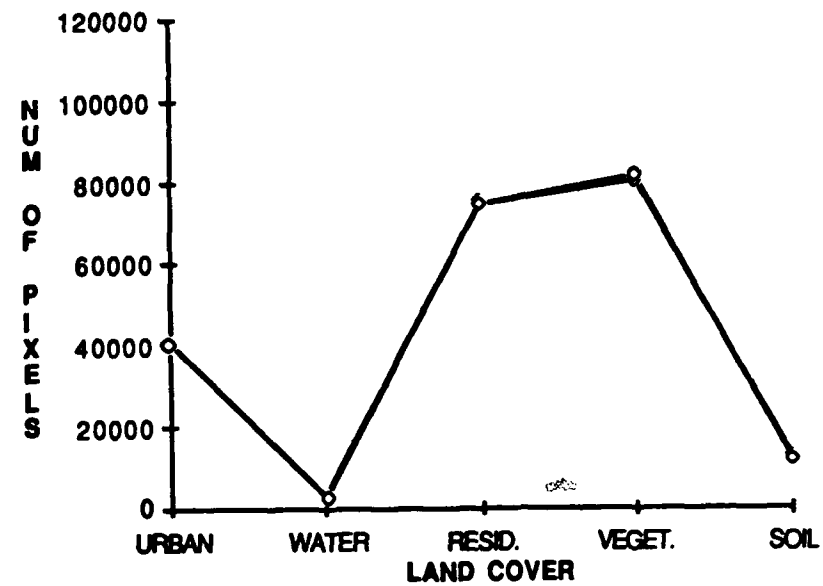


FIGURE II-4

Figure II-4: 1982-Mode Filter Comparison

Note: Two sets of data are plotted, but appear as one line in the graph.

After final processing of the classified images, the images were assessed for accuracy of the classification. This was done by sorting each image according to land cover class and having the computer randomly select 50 points from each class. These points were displayed to a monitor, one at a time using a 1:4 zoom factor to allow the experimenter to verify the location of the ground sample point and its actual land cover at the time the Landsat image was taken. The ground truth used for this process varied from high altitude color-infrared air photos taken as part of the National High Altitude Photography program to 35mm slides taken by the US Department of Agriculture as part of their crop monitoring plan. The ground truth for 1985 was fairly easy to obtain, but 1982 aerial photographs proved to be [illegible]sive. The ground truth used for 1982 was black and white high altitude photographs taken in 1981 and obtained from the U.S. Geological Survey's EROS Data Center. The year time difference of photography introduced some problems in identifying the vegetation and soil classes but was not much of a factor in locating the urban, residential and water classes. The classification accuracies determined by this process are contained in the Results section.

2.2.4. Creation of Difference Images.

After the images had been classified, the remaining step for change detection was a simple subtraction of the images. The DIRS Laboratory image processing software has the capability to subtract image planes allowing creation of the difference images. Before subtraction, a correction factor of 130 digital counts was added to the 1985 scene to raise all the final subtraction values above zero. Also the 1985 scene had some border areas removed that were not found in the registered 1982 scene. The resulting difference images contained pixel values of 130 in locations where no change had occurred between the classified scenes and values of other than 130 where the two classified images were different.

2.2.5. Change Detection.

This step was simplified by the earlier classification of the original images on the basis of assigned gray levels. The changes were categorized by their difference values as each change has its own unique gray level difference value resulting from the appropriate selection of classification values. These values are shown in table II-7 as a difference image matrix.

Table II-7

Difference Image Change Detection Matrix

Class in 1985:	Class in 1982: URB	WAT	RES	VEG	SOIL
URB	130	115	90	55	10
WAT	145	130	105	70	25
RES	170	155	130	95	50
VEG	205	190	165	130	85
SOIL	250	235	210	175	130

2.2.6. Change Detection Accuracy

A sampling plan similar to that described for the classified images was used where 50 random samples were selected in five categories of change. These five categories of change covered 92.2 percent of the 85-82(classified with 1982 class mean vectors) difference image and 87.9 percent of the 85-82(PIF) difference image as illustrated in table II-8. Each sample point was checked as to actual land cover class and to the type of change that occurred using the ground truth previously described. An error matrix similar to

table II-4 was created for each image and included in the results section.

Table II-8

Difference Image Changes Detected

(As A Percent of the Total Image)

82 - 85 Change	1985	1982
No Change	74.0%	65.9%
RES - URB	2.0%	4.3%
VEG - RES	8.3%	15.0%
RES - VEG	4.2%	1.3%
URB - RES	3.7%	1.4%
Total	92.2%	87.9%

2.3. Post-Classification Change Detection (One algorithm)

A comparison change detection was made from the uncorrected 1982 scene. This uncorrected scene was registered using the same bilinear interpolation resampling technique previously mentioned. The registered scene was then classified with the 1985 classification means shown in table II-3 to create the same five classes. This method was faster to

implement than the previous two classifier method, but made no corrections for the differences in atmosphere or illumination between the two scenes. The accuracy of the classification was verfied using 50 random samples per class against the same ground truth previously mentioned. The classified image was then subtracted from the 1985 classified scene to create the difference image. The difference image values found in table II-7 are applicable to this difference image also. The results of this portion had only limited success and are discussed later.

2.4. Gain and Offset Correction Technique (Using one algorithm)

The previous change detection method did not attempt to correct for known calibration differences between the two different sensors, Landsat 4 and Landsat 5, when comparing the two scenes. As a possible improvement to the accuracy of that method, This next method takes into account the known gain and offset differences between the calibration of the two sensors.

2.4.1. Gain and Offset Correction.

The method discussed earlier for correcting the gain and offset differences between sensors was used to improve the accuracy of the 1985 classifier on the 1982 imagery. As developed before in equation (10), the Digital Count of one sensor can be corrected for the

differing gain and offset of the other sensor:

$$DC_1 = 255 \times \frac{(Lmin_2 - Lmin_1)}{(Lmax_1 - Lmin_1)} + \frac{(Lmax_2 - Lmin_2)}{(Lmax_1 - Lmin_1)}(DC_2) \qquad (10)$$

The actual calculations for these two images are found in Appendix B, but the resulting slope and intercept corrections for each band in the 1982 image are shown in table II-9.

Table II-9

Gain and Offset Corrections from a 1982 to a 1985 TM Scene

BAND	GAIN	OFFSET
1	1.0413	-0.0332
3	1.1474	0.0372
4	1.0873	-0.0123
5	1.898	0.0
7	1.1803	0.0

2.4.2. Registration of the Two Scenes.

The corrected 1982 scene was then registered to the 1985 scene using the same registration method of bilinear interpolation with a resampling of the 1982 image. This image contained the same visual registration errors as the previous two registered images.

2.4.3. Land Cover Classification.

The class mean vectors (found in table II-3) developed using discriminant analysis and training areas identified on the 1985 image were used to classify the corrected image. The mode filter was applied to the image produced by the minimum distance to the mean classifier. The resulting classified image was randomly sampled (50 samples in each of five classes) and the samples were checked against the ground truth for classification accuracy. The classified image was then subtracted from the 1985 classified image to create a change detection difference image in the manner described before. The same values shown in table II-7 apply to this difference image.

2.5. PIF Normalization Technique (One classification algorithm)

The PIF normalization process applies a radiometric normalization to the two images. In actual application, "Statistical alysis of

segmented scene elements is used to generate band by band transformation functions."[3] This normalization overcomes differences due to illumination and atmosphere between the images, allowing the pixel digital counts to be more accurately compared. This PIF normalization allowed use of the class signatures developed for a 1985 scene to be used on a 1982 scene. This is a time and computer cost savings over the standard method of building a second class signature set (in the form of class means as in table II-5).

2.5.1. PIF Normalization of Original Subscenes.

The PIF normalization is built upon the temporal invariance of reflectance values for urban areas composed of asphalts and concretes which must first be separated from the remainder of the scene. The segmentation of asphalts and concretes can be performed by interactive thresholding of each subscene to allow identification of the urban features that act as PIFs. DIRS Laboratory staff have also developed a computer automated technique to perform this segmentation with repeatable results. The automated method was the technique used in this thesis.[33] These segmented urban features are then separated from the remaining pixels by use of a classification mask. This mask is applied to each of the spectral bands of interest leaving just the PIF areas in each band. The histogram for the PIF area in each day 1 scene is calculated and compared to the PIF area in the appropriate day 2 subscene. From these histograms a

transformation equation for each day 2 band is developed. These equations are used to transform the day 2 image to appear as if it had been imaged through the same atmosphere as the day 1 image. The transform was implemented using a "look up table" or Intensity transformation table (ITT). The image transformation data is summarized in table II-10.

Table II-10

Summary of PIF Normalization and Transformation Data for the 1982 TM Scene

	PIF Histogram Statistics				Linear Transformations	
	1982		1985		1982 to 1985	
Band	Mean	Std Dev	Mean	Std Dev	slope	intercept
1	107.5	8.4	129.7	17.9	2.13	-99.9
3	46.6	7.4	74.4	18.0	2.43	-38.6
4	42.2	7.2	69.9	16.4	2.29	-26.7
5	58.4	15.0	104.5	30.1	2.00	-12.3
7	34.16	9.7	63.9	21.5	2.21	-11.7

2.5.2. Registration of Scenes.

The normalized 1982 scenes were then registered to the unchanged 1985 scenes using the same bilinear interpolation technique discussed in the previous registrations.

2.5.3. Land Cover Classification + Difference Image Creation.

The same class signatures derived for the 1985 image (see table II-3) using a supervised training set, were used to classify the PIF normalized 1982 image. This classification was performed using the minimum distance to the class mean method run on the same LIPS software previously mentioned. The 3 x 3 Mode Filter was performed on the classified image with the results shown in figure II-5.

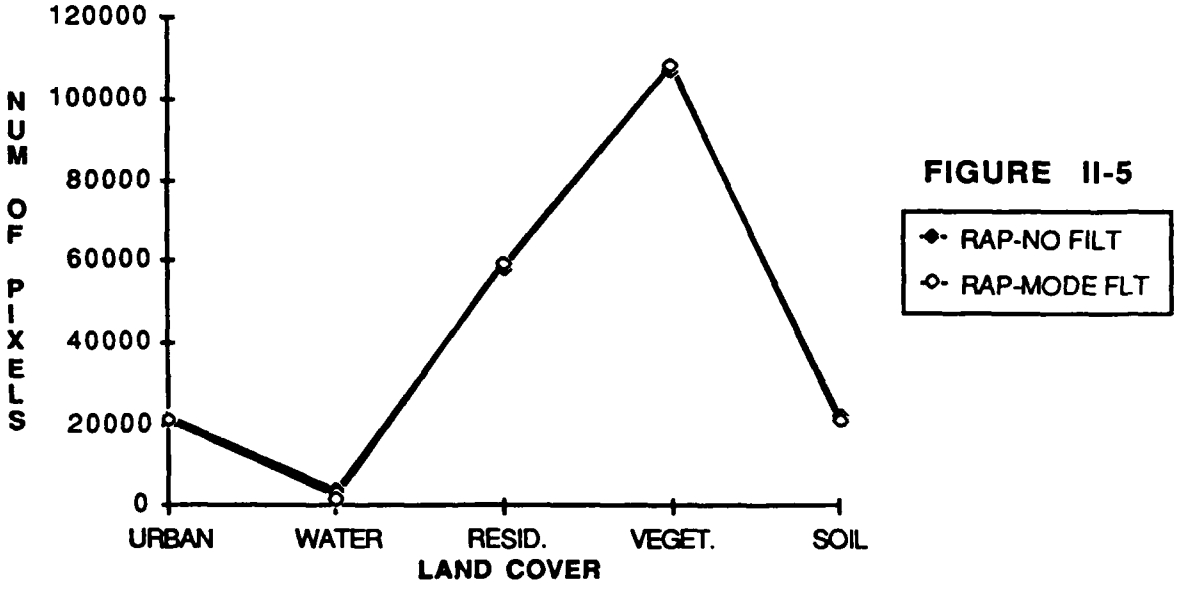


FIGURE II-5

Figure II-5: 1982 PIF-Mode Filter Comparison

Note: Two sets of data points are plotted, but appear as one line.

2.5.4. Classification Error Matrix.

An error matrix similar to the one shown in table II-4 was developed to quantify the resulting classification accuracy. The actual data will be discussed in the results section.

2.5.5. Creation of Difference Images.

Again the 1982 classified image was subtracted from the 1985 classified image to create a difference image that would allow easy identification of changes. This was accomplished by choosing the digital counts for each classification such that the results conformed to the matrix previously shown in table II-7.

2.5.6. Change Detection.

2.5.6.1. Identify changes detected in Difference Image.

Areas of no change showed a value of 130 digital counts for the difference image while those areas experiencing change showed some value other than zero. Again, the difference values were adjusted with a correction factor to eliminate negative values.

2.5.6.2. Accuracy Verification

A sampling plan similar to that described for the classified images was again used where 50 random samples were selected in five categories of change. These five categories of change covered 92.2 percent of the 85-82(classified with 82 class mean vectors) difference image and 87.9 percent of the 85-82 (PIF) difference image as illustrated in table II-8. Each sample point was checked as to actual land cover class and to the type of change that occurred using the ground truth previously described. An error matrix similar to II-4 was created for each image and included in the results section.

3.0. Results and Discussion.

This section is divided into two portions, one discussing the results of classification of the Landsat images from both the 1982 data and the 1985 data. The other portion discusses the change detection results and reasons for the changes observed. This format is used because the change detection accuracy using the post-classification method is directly tied to the accuracy of the classifiers implemented. Any classifier inaccuracy in either image is immediately translated into a wrong class of change detected in the difference image.

3.1. Classification Accuracy.

All of the image classifications were performed using a minimum distance to the class mean method which measures the distance from a given pixel vector of an image pixel (the vector is composed of the digital count values for bands 1, 3, 4, 5, and 7 at that particular image location) to each of the five class mean vectors. The pixel is assigned to the same class as the closest class mean vector. These class mean vectors were developed earlier using an iterative process of point selection and linear discriminant analysis on the original 1985 or resampled 1982 image. The training points selected were used to create the class mean vectors. The linear discriminant algorithm, minimum distance to the mean classifier, and the training point selection software are available in the LIPS software package contained on the

DIRS Laboratory VAX/Gould Deanza computer system.

3.1.1. 1985 Scene Classification Accuracy.

The training set developed for the 1985 image was tested on itself using the minimum distance to the mean method with the results reported in table II-4. The 97.5% overall performance rating is an estimate that 97.5% of the pixels in the training set were classified accurately into their own classes with the remainder showing up in other classes. These incorrectly identified pixels are pixels that were contained within the boundry of one training area but were composed of pixel vector values that were more closely identified with another class mean. An example of this kind of misclassified pixel was vegetated areas found within a residential setting where the amount of vegetation is heavier than that found in the areas used for the class mean vector. These pixels are correctly identified as vegetation by the classifier even though contained in a residential training set.

In an attempt to add spatial information to the classifier process, a standard deviation image was created by passing a 3x3 window over the image which deposited the standard deviation value for the nine pixels covered by the window at the center point of the window. Hsu (1978) has reported the ability of this spatial information to increase the accuracy of black and white aerial photographs. In this case, the

standard deviation image did not increasase the accuracy of the classifier, but actually confused the classifier. The 30m by 30m pixel size of the Landsat image did not allow enough resolution for the standard deviation to add accurate information to the classification, so this additional band of spatial information was not used in classifying any of the images.

The 1985 image was then classified using the minimum distance classifier and the class mean vectors developed from the training set. The resulting image is shown in figure III-1.

Figure III-1: 1985 Classified Rochester Image.

The accuracy of the classification process was then tested using the random sampling technique previously discussed. Fifty random samples from each class were selected and displayed on the computer monitor at a 1:4 zoom to allow identification of the pixels for location and class. The 1985 color-infrared aerial photographs obtained from the U. S. Geological Survey and the 1985 35mm color slides obtained from the U.S. Department of Agriculture were used as ground truth to determine the actual land cover classification. The verification results are recorded in matrix format in table III-1.

Table III-1

1985 Classification Accuracy Determined by

Ground Truth Comparison.

Classification in Image:	Classification Found In Ground Truth:					Percent
	Urban	Water	Resid.	Veget.	Soil	Correct
Urban	44	0	3	1	2	88
Water	3	47	0	0	0	94
Residential	7	1	38	2	2	76
Vegetation	1	0	3	45	1	90
Soil	10	0	1	5	34	68

OVERALL PERFOMANCE = 83%

Given the sample size of 50 samples and a desired degree of confidence (say 95 percent confidence) the appropriate confidence interval for each of the classes can be calculated for the above accuracy values using the binomial distribution. In this case an 88 percent correct value for the urban class results in a lower confidence limit of 76.2 percent and an upper limit of 94.4 percent with a 95 percent confidence based on the method of Hord Brooner (1976).[22] A similar interval can be calculated for each class based on the confidence desired.

The percent of correctly identified pixels is not as high as the result of the training set performance, but the actual scene classification is expected to be degraded due to factors such as mixed pixels and class boundry areas that were avoided in the training set by picking the training set locations in the middle of known class areas where there was little chance of these degrading effects occuring. Particularly clear from the results is the difficulty that the Soil classification had in discriminating between soil and some bright urban areas. This was especially evident along the main railroad line leading into Rochester from the east where a number of the large storage buildings had highly reflective roofing materials which were interpreted by the classifier as soil. The Soil classifier also had problems with some of the light concerete areas, especially near the airport. These high reflectance surfaces also appeared to the classifier to be closer to the Soil class mean vector than the less

highly reflective Urban mean vector. Part of this problem might be solved by using more classes and giving these more highly reflective surfaces a separate class to avoid the confusion with Soil.

The Soil classifier had difficulties with the vegetative side of the classification scale as well. These misclassifications occured in areas of light vegetative growth where the supporting soil was visible through the vegetation. This is a more difficult problem to counteract as the type of vegetation also has an impact, along with the possible shadow conditions caused by the vegetation in relation to the illumination (in this case, primarily direct sunlight). Sparsely growing grass casts very little shadow and hence looks more like soil, while rows of leafy bean plants may, at the right sun angle or sensor look angle, appear as thick vegetation. These same leafy plants could also return a very high reflectance value if, for instance, the sun angle approaches 90 degrees and there is little shade preventing the soil between the rows from causing a high reflectance return. Here also, more classes in the classifier would allow a finer distinction between the types of vegetation and perhaps avoid some of the misclassification of vegetation as soil.

The Residential class was the next area where the minimum distance to the mean classifier had difficulty in separating different classes. Some of this confusion was caused by the author's definition of residential areas that included apartment complexes which

contained large buildings along with significant parking areas. These tended to have reflectance vectors closer to the Urban mean class vector than the more vegetative influenced Residential areas. Likewise, there were some light industrial areas that were certainly urban features, but that also had large amounts of vegetation around the buildings. This vegetation made the reflectance values appear closer to the Residential mean vector. The mixed pixel effect was also a factor as an Urban area next to a Vegetation area did not usually coincide with the "footprint" of the sensor's pixels on the ground and in this case the pixel would see the mixed reflectance of the Urban area and the Vegetative area, which together appear closer to the mean vector values for the Residential class than either of the other two classes.

The Water classification was fairly accurate at 94%, but there were some misclassifications with the Urban class. This is due to both the mixed pixel effect and to the low reflectance values received from some urban features such as ashphalts.

3.1.2. 1982 Scene Classification Accuracy (1982 Mean Vectors).

The registered 1982 image was also sampled for construction of a training set to develop class mean vectors that would correctly classify the image pixels. The results of this process are the class

means shown in table II-5 and were tested using the minimum distance to the mean method. The confusion/error matrix is found in table II-6. The entire 1982 image was then classified using the CLAMID software with the classified image shown in figure III-2.

Figure III-2: Classified 1982 Rochester Image (Using 1982 Class Mean Vectors)

The accuracy of this classification process was tested by verifying the ground truth of 50 samples in each class with the results recorded in table III-2. Unfortunately, ground truth for the 1982 time period was not available, so black and white aerial photos taken in 1981 were used for the primary ground truth with some

reference to the previously mentioned 1985 ground truth for further location verification. This may have introduced some error into the accuracy assessment, but allowance for possible change from 1981 to 1982 was necessary in only a small percentage of the sample points, so this error was assumed to be negligable.

Table III-2

1982 Classification Accuracy Determined by Ground Truth Comparison.

Classification in Image:	Classification Found In Ground Truth:					Percent
	Urban	Water	Resid.	Veget.	Soil	Correct
Urban	46	0	4	0	0	92
Water	11	39	0	0	0	78
Residential	3	3	43	1	0	86
Vegetation	2	0	6	40	2	80
Soil	13	1	2	5	29	58

Overall Performance = 83%

One factor that affected the classification accuracy here was the atmospheric haze that is visually apparent in the 1982 image. This haze had the effect of attenuating the reflectances available for

classification. This effect is apparent when the class mean vectors for the 1985 classifier (see Figure II-3) are compared to those for the 1982 classifier (see Figure II-4). The compressed range of the digital counts for the 1982 class mean vectors is quickly seen from these graphs.

The Soil class was again the most inaccurate of the classifications as some urban areas continued to be confused as soil due to the high reflectance from their highly reflective roofing materials, like metal and crushed white stone. Also some concretes were of a high enough reflectance value to be mistaken for soil. Some sparsely vegetated areas were also mistaken for soil.

The Water classification again had difficulty in separating some of the urban features with very low reflectances, but the accuracy should actually be higher as some of the points picked by the compter were along an edge of the image that was classified as water, but should be actually considered a border area, which could be removed from the image.

The Vegetation classification misclassified some residental areas due to the large amounts of vegetation present among the actual structures. The Residential classification suffered from some mixed pixel effects as it mistook mixed water and vegetation pixels as residential areas. The Urban class had only small confusion with

some residential areas which were along boundries between the two classes.

3.1.3. 1982 Image Classified with 1985 Class Mean Vectors

To test the accuracy of using a one classifier method, the 1982 image was classified using the class mean vectors developed for the 1985 image. This method was certainly less time consuming as only one classifier training set was nee d to classify both images. The problem with this method was that it did not take into account the differences in sensor, illumination, and atmosphere between the two images. These neglected differences affected the ability of this post- classification method to accurately classify the 1982 image and then to accurately identify the changes. The classified image is shown in Figure III-3. The classified image was sampled with 50 samples in each of the five classes in the manner dicscussed previously. The results are shown below in table III-3.

Figure III-3: Classified 1982 Image Using 1985 Class Mean Vectors

Table III-3: 1982 Image Classified using 1985 Class Mean Vector. Classification Accuracy Determined by Ground Truth Comparison.

Classification in Image:	Classification Found In Ground Truth:					Percent
	Urban	Water	Resid.	Veget.	Soil	Correct
Urban	42	1	3	1	3	84
Water	18	12	20	0	0	24
Residential	4	1	25	20	0	50
Vegetation	0	0	1	49	0	98
Soil	41	0	0	0	9	18

Overall Performance = 52%

The results shown above indicate the inaccuracies produced by this simple method. The high accuracy attained by the Vegetation class does not reflect the small number of actual pixels that were identified as being Vegetation (only 290). Most of these Vegetation pixels were actually located on golf courses. The vast majority of the image pixels (166,738) were identified as Residential with only a 50 percent accuracy. The percentage of pixels in each classification is compared to the 1982 image classified with its own training set and the 1985 image classified with its own training set in figure III-4 to help illustrate the different results between the methods.

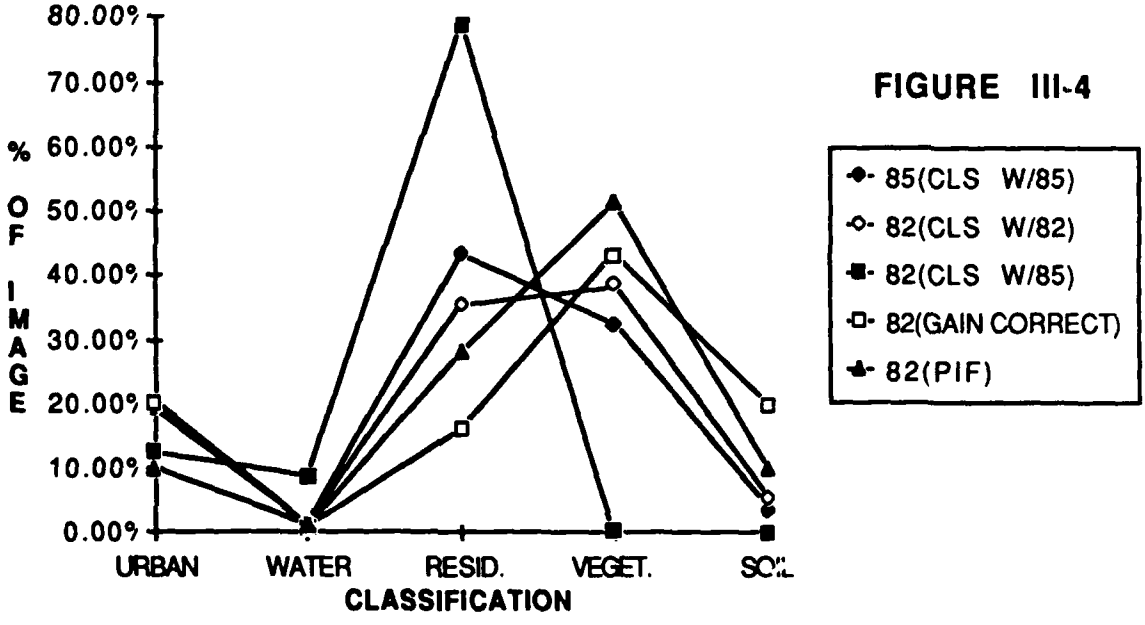


FIGURE III-4

Figure III-4: Comparsion of Classifications of All Methods

3.1.4. 1982 Gain and Offset Corrected Scene Classified with 1985 Class Mean Vectors.

The Gain and Offset correction discussed previously was applied to the 1982 image in an attempt to account for the different sensor calibration levels between Landsat 4 (at the time of the image) and Landsat 5. These corrections still did not account for different atmospheric and illumination conditions between the two images, but the corrections were simple to implement and might have provided an increase in the accuracy of the classification. Again, the corrected image was classified using the class mean vectors derived from the 1985 image. The classified image is shown in figure III-5. The results of the 50 sample accuracy verification performed for each classification are shown in table III-4.

Figure III-5: 1982 Gain and Offset Corrected Image Classified using 1985 Class Mean Vectors

Table III-4

1982 Gain and Offset Corrected Image Classified using 1985 Class Mean Vectors, Classification Accuracy Determined by Ground Truth Comparison.

Classification in Image:	Classification Found In Ground Truth:					Percent
	Urban	Water	Resid.	Veget.	Soil	Correct
Urban	28	0	21	1	0	56
Water	4	33	1	10	0	66
Residential	5	1	38	6	0	76
Vegetation	5	3	12	28	2	56
Soil	23	0	6	13	8	16

Overall Performance = 51%

The results of this classification method are as equally inaccurate and inconclusive as the previous method. As indicated by figure III-4, the proportion of the image assigned to the various classes is grossly different from the 1982 image classified with the 1982 training set which must be used for a comparison.

3.1.5. 1982 PIF Normalized Scene Classified With 1985 Class Mean Vectors.

The PIF normalization previously discussed was performed on the 1982 image to account for the differences in sensor calibration levels, illumination and atmosphere that had occured between the images. The normalized image was then classified using the 1985 class mean vectors shown in table II-3. The resulting classified image is shown in figure III-6. The results of the accuracy verification using 50 samples from each of the five classes is shown in table III-5.

Figure III-6: 1982 PIF Normalized Image Classified using 1985 Class Mean Vectors

Table III-5

1982 PIF Normalized Image Classified using

1985 Class Mean Vectors,

Classification Accuracy Determined by Ground Truth Comparison.

Classification in Image:	Classification Found In Ground Truth:					Percent
	Urban	Water	Resid.	Veget.	Soil	Correct
Urban	47	0	3	0	0	94
Water	3	47	0	0	0	94
Residential	7	1	38	3	1	76
Vegetation	1	0	9	39	1	78
Soil	12	0	5	7	26	52

Overall Performance = 77%

The Soil classification was again the worst performing section of the classification, with Urban areas still providing the majority of the false classifications due to high reflectance roofing material and concrete surfaces which resembled the reflectances from soil. There continued to be some confusion with sparsely vegetated areas as well.

The Vegetation classification had difficulty eliminating some

residential areas from its class with large sections of vegetation continuing to be a problem of classification in Residential areas. The Residential classifier had its largest difficulties with Urban areas where the mixed pixel effect caused by adjoining urban and vegetation areas was a factor. Also some Vegetation areas were mistakenly classified as Residential areas.

The high accuracy of the Urban classification was expected as the PIF Normalization is based on the transformation of urban like features which act as the PIF. The confusion observed with the residential class is again caused by large apartment complexes which were classified as residential by the author, but that have reflectance vectors similar to urban conditions due to the large buildings with large parking areas. The Water classification again had some small confusion with low reflectance urban areas, but overall performed as accurately as the Urban classification.

3.2. Change Detection Accuracy.

Due to the large inaccuracies found during the classification process, only the 1982 image classified with its own 1982 class mean vectors and the 1982 PIF normalized image classified with the 1985 class mean vectors were selected to continue to the change detection process with the 1985 classified image.

A constant value of 130 digital counts was added to the 1985 classified image shown in figure II-4 to allow subtraction without creating negative numbers. First, the 1982 image classified with its own 1982 class mean vectors was subtracted from this artifically increased 1985 image with the difference image shown in figure III-7. Then, the 1982 PIF normalized image classified with the 1985 class mean vectors was subtracted from the same 1985 image with the difference image shown in figure III-8.

Figure III-7: 1985-1982 (Two Classifiers) Difference Image.

Figure III-8: 1985-82 PIF (One Classifier) Difference Image

The digital counts of the difference image are determined by the starting values of the classified image. If no increase had been made to the values of the 1985 image, the difference image would contain zero values at any areas where the same class existed in each original image. With the addition of 130 digital counts to the entire 1985 classified image, these zero areas were represented as digital counts of 130. This increase is the result represented in the difference images. Through use of the classification values shown in table III-6, the difference image contains the changes between classifications according to the actual digital count of the difference

images. The key to interpreting these digital counts is shown in table III-7.

Table III-6

Classified Image Digital Count Assignments

Class	Digital Count
Urban	5
Water	20
Residential	45
Vegetation	80
Soil	125

Table III-7

Key To Difference Image Values

Class in 1985:	Class in 1982: Urban	Water	Resid.	Veget.	Soil
Urban	130	115	90	55	10
Water	145	130	105	70	25
Residential	170	155	130	95	50
Vegetation	205	190	165	130	85
Soil	250	235	210	175	130

The difference images are largely composed of just a few of the change classes with the "no change" value of 130 covering most of the images. The actual number of pixels in each class is shown in Table III-8 and depicted graphically in figures III-9 and III-10 as Histograms.

Table III-8

Difference Images' Histogram Data

Change	85-82 (Two Classifier) DC	# Pix	%	85-82 (PIF) # Pix	%
Soil-Urb	10	2628	1.00%	7423	2.83%
Soil-Wat	25	19	0.01%	70	0.03%
Soil-Res	50	2359	0.90%	5149	1.96%
Veg-Urb	55	3438	1.31%	5672	2.16%
Veg-Wat	70	93	0.04%	551	0.21%
Soil-Veg	85	3847	1.47%	5002	1.91%
Res-Urb	90	5299	2.02%	11285	4.30%
Veg-Res	95	21758	8.30%	39303	14.99%
Res-Wat	105	843	0.32%	1031	0.39%
Wat-Urb	115	447	0.17%	273	0.10%
No Change	130	193935	73.99%	172701	65.88%
Urb-Wat	145	203	0.08%	250	0.10%
Wat-Res	155	470	0.18%	715	0.27%
Res-Veg	165	11015	4.20%	3412	1.30%
Urb-Res	170	9574	3.65%	3641	1.39%
Veg-Soil	175	3781	1.44%	3703	1.41%
Wat-Veg	190	95	0.04%	656	0.25%
Urb-Veg	205	768	0.29%	382	0.15%
Res-Soil	210	500	0.19%	415	0.16%
Wat-Soil	235	24	0.01%	81	0.03%
Urb-Soil	250	1010	0.39%	429	0.16%

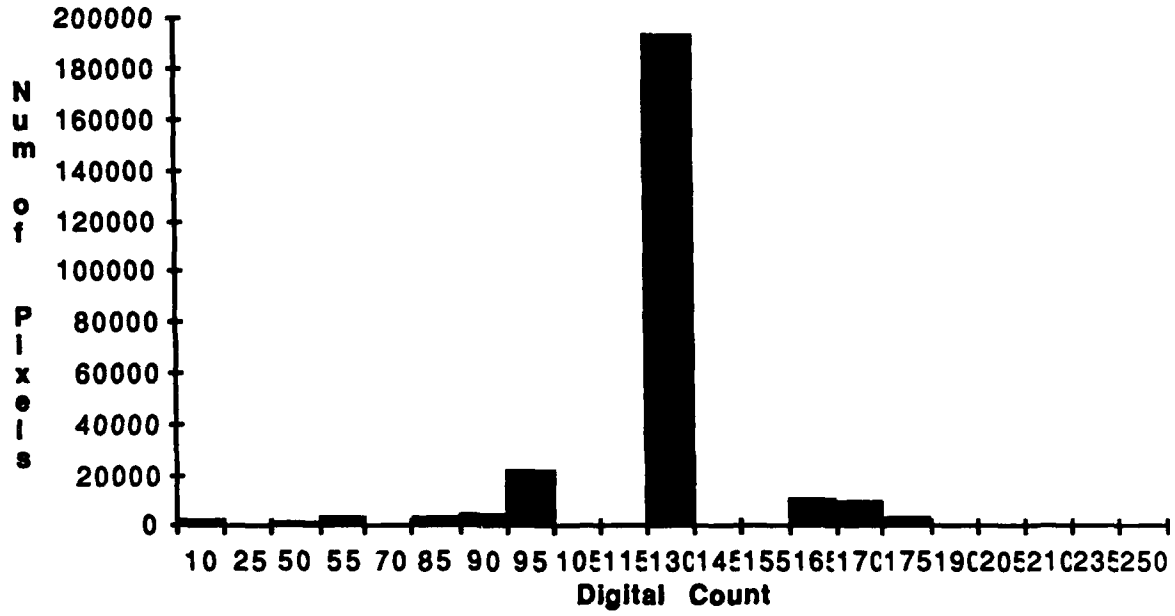


Figure III-9: 85-82 Difference Image Histogram

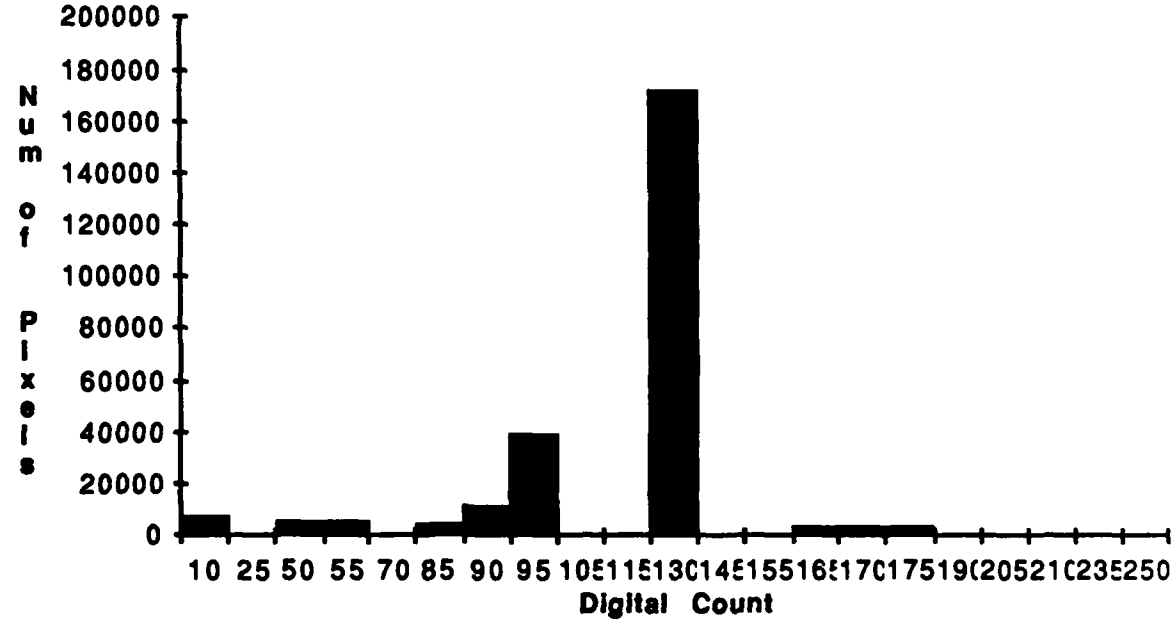


Figure III-10: 85-82 PIF Difference Image Histogram

These figures illustrate the large portion of the image defined into a small number of classes. In fact, as table III-9 shows, 92 percent of the two classifier difference image and 88 percent of the PIF difference image are defined by the five largest classes of the 85-82 two classifier difference image.

Table III-9

Five Largest Classes of Two Classifier Difference Images

Class Description	Digital Count	85-82(2class)	85-82(PIF)
Resid to Urban	90	2.0%	4.3%
Veget to Resid	95	8.3%	15.0%
No Change	130	74.0%	65.9%
Resid to Veget	165	4.2%	1.3%
Urban to Resid	170	3.7%	1.4%
	Total	92.2%	87.9%

For the two classifier image, 50 random samples from each of these five change classes were choosen and then verified against the ground truth used in the previous sections. The results of this verification are shown in table III-10. For the No Change class, the

accuracy was very high at 98 percent with the only change missed being a soil to vegetation change. The remaining classes did not perform as well. Most of the randomly sampled points actually experienced no change in the ground truth and had indicated change because one of the classifiers had misclassified a particular pixel. Many of these misclassifications occured near border areas where the probabilities of mixed pixels are increased. The Vegetation to Residential change class had actually detected the appropriate change but, this was with only 3 out of the 50 samples in the change class.

Table III-10

1985-1982 (Two Classifier) Difference Image Accuracy for Selected Change Detection Classes

	Change Class in Ground Truth:						
Change Class in Dif. Img.	No Change	Veg-Urb	Veg-Res	Soil-Veg	Soil-Urb	Soil-Res	Percent Correct
Resid to Urban	48	2	0	0	0	0	0%
Veget to Resid	45	2	3	0	0	0	6%
No Change	49	0	0	1	0	0	98%
Resid to Veget	50	0	0	0	0	0	0%
Urban to Resid	47	0	0	0	2	1	0%

The results were very similar in the PIF difference image. The No Change class displayed 100 percent accuracy for the 50 samples selected. Unfortunately, all the other change classes also experienced a 98 percent or greater no change condition in the actual ground truth. In this image also, many of the mistaken classifications occurred along land cover class borders in one of the classified images. The specific results are shown in table III-11.

Table III-11

1985-1982 (PIF) Difference Image Accuracy for Selected Change Detection Classes

	Change Class in Ground Truth:						
Change Class in Dif. Img.	No Change	Veg-Urb	Veg-Res	Soil-Veg	Soil-Urb	Soil-Res	Percent Correct
Resid to Urban	49	0	0	0	0	1	0%
Veget to Resid	49	0	0	1	0	0	0%
No Change	50	0	0	0	0	0	100%
Resid to Veget	50	0	0	0	0	0	0%
Urban to Resid	49	0	0	0	1	0	2%

4.0. Conclusions and Recommendations.

The concept of using Pseudo-Invariant Feauture normalization to allow simpler classification of Landsat images has been demonstrated on a scene composed of a mixture of urban and rural features. This was demonstrated using standard image classification techniques on Landsat images that are readily available. The overall classification accuracy results are summarized below in table IV-1.

Table IV-1: Overall Classification Accuracy Results

Method	Overall Accuracy
1985 Classified w/ 85 Class Mean Vectors	83%
1982 Classified w/ 82 Class Mean Vectors	83%
1982 PIF Normalized - Classified w/ 85 Class Mean Vectors	77%
1982 Classified w/ 85 Class Mean Vectors	52%
1982 Gain and Offset Corrected - Classified w/ 85 Class Mean Vectors	51%

The overall accuracy percentage was calculated by weighting the raw

accuracy percentages reported earlier by the number of pixels of the same class in that image for the different classification methods. The calculations are shown in Appendix C. These results show that the PIF Normalization allows classification of two temporally different images using the same class signatures with accuracies comparable to a more lengthy and time consuming standard method of using two different sets of class signatures.

The classification process was undertaken using standard classification techniques with an emphasis on minimizing the processing required to obtain an accurate classification. The PIF normalization allowed classification with results comparable to the more lengthy two classifier method, which involves creating supervised training sets for two images. The ability to remove man from part of the processing chain as demonstrated in the PIF normalization method is attractive in this time of high labor costs.

The PIF normalization could be used in the future to allow creation of a time-independent classification process where all the images are normalized to one standard scene. The images would then be classified using the standard scene classification algorithms. Since these standard scene classification algorithms will be used over and over, more time can be devoted to increasing the accuracy of this classifier. This process ultimately could lead to a set of time-indepedent classification techniques that would be applied to any appropriate PIF normalized image.

Two other techniques were attempted in order to demonstrate the classification accuracies achievable using even less complicated processing schemes. First, a classification of both images using the same class mean vectors without any scene correction was demonstrated and showed that the accuracies achievable were not comparable to the two more complicated classification methods. The uncompensated differences in atmosphere, sensor and illumination caused more errors than either of the two previous classification methods and was eliminated from the change detection demonstration because of the inaccuracies involved. A second technique attempted to account for some of the sensor differences by using a gain and offset correction, but did not attempt to account for atmospheric and illumination differences. The classification using the same class mean vectors also proved to be very inaccurate and this classified image was also excluded from the change detection demonstration.

The effect of image resolution on classifier accuracy was mentioned as a possible cause of inaccuracy in the classified image. This occurs as a result of the many pixels whose reflectance value is actually composed of the combined reflectances of more than one target or classification on the ground. The effect of these mixed pixels could be decreased by increasing the number of pixels used to image the same ground surface area. With more homogeneous reflectances the accuracy of the classifier should increase as the number of pixels falling into categories between class mean vectors will decrease. An increase in resolution might allow the use of spatial information, such as a standard deviation image, to increase the accuracy of

the classification. Also, using more classes in the classification process may allow the classifier to more accurately place pixels in the appropriate class rather than choosing the wrong class based on a limited choice and a mixed reflectance return.

As previously mentioned, the key to post-classification change detection is the accuracy of the classifications of the original images. The linear discriminant analysis method used to choose the class mean vectors and the minimum distance classifier are both standard image processing techniques and their results may be improved upon by other classification techniques such as a gaussian maximum likelihood classifier. In particular the ECHO technique and the "Tassel Cap" technique seem to offer possibilities for improvement.[11]

For change detection, the methods used here served as a proof of concept, but much work on the classification issue remains to be accomplished. Given the inaccurate results observed from most of the classes in the change detection images, the No Change class displayed the highest accuracies. It also accounted for over 65% of the PIF normalized difference image. If this large portion of the image can be excluded from a search for changes, the remaining change areas can be used as cues for where to begin looking for change between the two images. This is certainly an improvement from manually comparing the entire two images in a search for changes. A large first step has been taken in limiting the comparison that needs to be done. With further analysis, it may be determined which of the

change classes are the most likely indicators of change and the search started with them. These possibilities for change detection cueing are largely unexplored. Also the difference image is in a format that is easily adaptable to an automated search of the changes.

A different method of change detection which does not involve classification of the images before creating the difference images should also be explored. The development of the PIF normalization technique now allows a direct subtraction of the images with any areas of no change indicated by a zero value in the difference image. The changes that did occur would cause something other than a zero value in the difference image. A method of classifying the change areas would need to be developed, but it may be as simple as using some of the existing classification techniques demonstrated in this thesis. This technique offers many possibilities and at least avoids the misclassification problems until after the changes have been detected.

In conclusion, the PIF Normalization Technique has proved its worth as a very powerful tool for land cover classification. It has allowed the extension of class signatures to temporally different images. The amount of processing time to be saved remains to be quantified, but the removal of a second class signature set is a big step towards simplifying the multi-temporal classification process. This technique holds promise for future benefits as it normalizes images to account for differing atmospheres and illumination between images of the same ground scene.

5.0. REFERENCES

1. D.S. Simonett, Ed., Manual of Remote Sensing, 2nd Ed., Vol. 1 , American Society of Photogrammetry, Falls Church, VA, 1983.

2. T. M. Lillesand and R.W. Kiefer, Remote Sensing and Image Interpretation, John Wiley & Sons, New York, (1978).

3. W.J. Volchok and J.R. Schott, "Scene to Scene Radiometric Normalization of the Reflected bands of the Landsat Thematic Mapper", Proceedings of the S.P.I.E. Symposium, "Optical and Opto-Electronic Applied Sciences and Engineering, Innsbruck, Austria, April 1986.

4. R. A. Weismiller, S. J. Kristof, D. K. Scholz, P. E. Anuta, and S. A. Momin, "Change Detection in Coastal Zone Environments", Photogrammetric Engineering and Remote Sensing, vol. 43, no. 12, 1533-1539, (1977).

5. R.A. Schowengerdt, Techniques for Image Processing and Classification in Remote Sensing, Academic Press, New York, NY, 1983.

6. D. L. Toll, "Landsat-4 Thematic Mapper Scene Characteristics of a Suburban and a Rural Area", Photogrammetric Engineering and Remote Sensing, vol. 51, no. 9, 1471-1482, (1985).

7. J.C. Gervin, Y.C. Lu, R.L. Gauthier, J.R. Miller, and R.R. Irish, "The Effect of Thematic Mapper Spectral Properties on Land Cover Mapping for Hydrologic Modeling", Paper presented at the Nineteenth International Symposium on Remote Sensing of Environment, Ann Arbor, MI, October, 1985.

8. C. Sheffield, "Selecting Band Combinations from Multispectral Data", Photogrammetric Engineering and Remote Sensing, vol. 51, no. 6, 681-687, (1985).

9. K.G. Dean, Y. Kodama, and G. Wendler, "Comparison of Leaf and Canopy Reflectance of Subarctic Forests", Photogrammetric Engineering and Remote Sensing, vol. 52, no. 6, 809-811, (1986).

10. J. C. Price, "The Contribution of thermal Data in Landsat Multispectral Classification", Photogrammetric Engineering and Remote Sensing, vol. 47, no. 2, 229-236, (1981).

11. R. L. Kettig and D. A. Landgrebe, "Classification of Multispectral Image Data by Extraction and Classification of Homogeneous Objects", IEEE Transactions on Geoscience Electronics, vol. GE-14, no. 1, 19-26 (1976).

12. M.J. Jackson, P. Carter, T. F. Smith, and W.G. Gardener, "Urban Land Mapping from Remotely Sensed Data", Photogrammetric Engineering and Remote Sensing, vol. 46, no. 8, 1041-1050, (1980).

13. C. H. Tom and L. D. Miller, "An Automated Land-Use Mapping Comparison of the Bayesian Maximum Likelihood and Linear Discriminant Analysis Algorithms", Photogrammetric Engineering and Remote Sensing, vol. 50, no. 2, 193-207, (1984).

14. S. Y. Hsu, "Texture-Tone Analysis for Automated Land-Use Mapping", Photogrammetric Engineering and Remote Sensing, vol. 44, no. 11, 1393-1404, (1978).

15. P.H. Swain, S. B. Vardeman, and J. C. Tilton, "Contextual Classification of Multispectral Image Data", Pattern Recognition, vol. 13, no. 6, 429-441, (1981).

16. A. R. Huete, "Separation of Soil-Plant Spectral Mixtures by Factor Analysis", Remote Sensing of Environment, vol. 19, 237-251, (1986).

17. J. R. Carr, C. E. Glass, and R. A. Schowengerdt, "Signature Extension Versus Retraining for Multispectral Classification of Surface Mines in Arid Regions", Photogrammetric Engineering and Remote Sensing, vol. 49 no. 8, 1193-1199, (1983).

18. P.P. Sanchez and S.A. Taheri, "A Data Structure With Applications to Remote Detection of Environmental Change", Proceedings of the Nineteenth International Symposium on Remote Sensing of Environment, Ann Arbor, MI, October, (1985).

19. J.L. van Genderen, and B.F. Lock, "Testing Land-Use Map Accuracy", Photogrammetric Engineering and Remote Sensing, vol. 43, no. 9, 1135-1137, (1977).

20. G. H. Rosenfield, K. Fitzpatrick-Lins,and H. S. Ling, "Sampling for Thematic Map Accuracy Testing", Photogrammetric Engineering and Remote Sensing, vol. 48, no. 1, 131-137, (1982).

21. W. J. Todd, D. G. Gehring, and J. F. Haman, Landsat Wildland Mapping Accuracy, Photogrammetric Engineering and Remote Sensing, vol. 46, no. 4, 509-520, (1980).

22. R.M. Hord and W. Brooner, "Land-Use Map Accuracy Criteria," Photogrammetric Engineering and Remote Sensing, vol. 42, no. 5, 671-677, (1976).

23. A.M. Hay, "Sampling Designs to Test Land-Use Map Accuracy," Photogrammetric Engineering and Remote Sensing, vol. 45, no. 4, 529-533, (1979).

24. P. O. Adeniyi, "Digital Analysis of Multitemporal Landsat Data for Land-Use/Land-Cover Classification in a Semi-Arid Area of Nigeria", Photogrammetric Engineering and Remote Sensing, vol. 51, no. 11, pp1761-1774, (1985).

25. D.J. Gerson and L.K. Fehrenbach, "Temporal Image Normalization", Defense Mapping Agency, Final Report, January, 1983.

26. K. R. Piech, D. W. Gaucher, J. R. Schott, and P. G. Smith, "Terrain Classification using Color Imagery", Photogrammetric Engineering and Remote Sensing, vol. 43, no. 4, 507-513, (1977).

27. D.D. Egbert and F.T. Ulaby, "Effects of Angles on Reflectivity", Photogrammetric Engineering , vol. 38, 556-564, (1972).

28. G.C. Grogan, "A Model to Predict the Reflectance from a Concrete Surface as a function of the Sun-Object-Image Angular Relationship", M.S. Thesis, Rochester Institute of Technology, Rochester, NY, June, (1983).

29. T. H. Lo, F. L. Scarpace, and T.M. Lillesand, "Use of Multitemporal Spectral Profiles in Agricultural Land-Cover Classification", Photogrammetric Engineering and Remote Sensing, vol. 52, no. 4, 535-544, (1986).

30. B.L. Markham and J.L. Barker, "Landsat MSS and TM Post-Calibration Dynamic Ranges, Exoatmospheric Reflectances and At-Satellite Temperatures", Landsat Technical Notes, August, 3-8, (1986).

31. S.C. Freden and F. Gordon Jr., in Manual of Remote Sensing, 2nd Ed., D.S. Simonett, Ed., vol 1, American Society of Photogrammetry, Falls Church, VA, 1983.

32. W.J. Volchok, "A Study of Multispectral Temporal Scene Normalization using Pseudo-Invariant Features Applied to Landsat Thematic Mapper Imagery", Masters Thesis for the Master of Science Degree in Imaging Science at R.I.T., October,1985.

33. C.Salvaggio, "Automated Segmentation of Urban Features From Landsat Thematic Mapper Imagery For Use inPseudoinvariant Feature Temporal Image Normalization", M.S. Thesis, Center for Imaging Science, Rochster Institute of Technology, Rochester, NewYork, May 1987.

6.0. Appendices.

Appendix A

Control Points and Linear Regression to Develop

Transformation Equations from 1982 Image to 1985 Image Coordinates

Obs.	1982 - X	Residual X	1982 - Y	Residual Y
1	167	0.390	319	-0.027
2	107	-0.166	302	0.699
3	98	-0.160	270	-0.344
4	91	-0.003	133	-0.501
5	185	0.496	88	0.791
6	200	-0.379	54	-0.277
7	302	0.156	83	-0.220
8	493	0.162	235	-0.019
9	451	-0.077	287	-0.604
10	475	0.235	363	0.199
11	442	-0.182	477	0.285
12	316	-0.178	487	-0.187
13	202	-0.069	447	-0.120
14	303	0.437	440	-0.047
15	343	-0.058	371	-0.032
16	386	-0.312	237	-0.219
17	249	-0.212	225	0.122
18	423	-0.114	187	0.479

Sum of Squared Residuals X = 1.1243

Error in Pixels X = 0.2499

Sum of Squared Residuals Y = 2.4408

Error in Pixels Y = 0.3682

Where the error in pixels = $\sqrt{[(\Sigma \text{ Residuals})^2 / N]}$

N = 18 samples

Appendix B

Gain and Offset Correction from Landsat 4 to Landsat 5

After Markham and Barker (1986)[30]

Band	82-L_{min}	82-L_{max}	85-L_{min}	85-L_{max}
1	-0.152	15.842	-0.15	15.21
3	-0.117	23.463	-0.12	20.43
4	-0.151	22.432	-0.15	20.62
5	-0.037	3.242	-0.037	2.719
7	-0.015	1.700	-0.015	1.438

$$\text{Transform 82} = 255 \times \frac{(Lmin_{82} - Lmin_{85})}{(Lmax_{85} - Lmin_{85})} + \frac{(Lmax_{82} - Lmin_{82})}{(Lmax_{85} - Lmin_{85})}(DC_{82})$$

$$\text{Transform slope (m)} = \frac{(Lmax_{82} - Lmin_{82})}{(Lmax_{85} - Lmin_{85})}$$

$$\text{Transform offset (b)} = \frac{(Lmin_{82} - Lmin_{85})}{(Lmax_{85} - Lmin_{85})} \times 255$$

Band	m*	b*
1	1.0413	-0.0332
3	1.1474	0.0372
4	1.0837	-0.0123
5	1.1898	0.0
7	1.1803	0.0

* Used to transform 1982 image to 1985 calibration values using:

$$\text{transformed 82} = m\,(82\ DC) + b$$

Appendix C

Overall Classification Accuracy Performance Calculations

	ACCURACY	# PIXELS	% OF IMAGE	WEIGHTED ACCUR	
	85(CLS W/85)				
URB	88	40803	19.25%	16.94%	
WAT	94	2873	1.36%	1.27%	
RES	76	91952	43.38%	32.97%	
VEG	90	68759	32.44%	29.20%	
SOIL	68	7562	3.57%	2.43%	
		211949	100.00%	82.81%	OVERALL PERF
	82(CLS W/82)				
URB	92	40515	19.16%	17.62%	
WAT	78	2749	1.30%	1.01%	
RES	86	75020	35.47%	30.51%	
VEG	80	81676	38.62%	30.90%	
SOIL	58	11523	5.45%	3.16%	
		211483	100.00%	83.20%	OVERALL PERF
	82(CLS W/85)				
URB	84	26360	12.44%	10.45%	
WAT	24	18441	8.70%	2.09%	
RES	50	166738	78.67%	39.33%	
VEG	98	290	0.14%	0.13%	
SOIL	18	120	0.06%	0.01%	
		211949	100.00%	52.01%	OVERALL PERF

	ACCURACY	# PIXELS	% OF IMAGE	WEIGHTED ACCUR	
	82(Gain+Offset)				
URB	56	42352	20.15%	11.28%	
WAT	66	1763	0.84%	0.55%	
RES	76	33937	16.15%	12.27%	
VEG	56	90718	43.16%	24.17%	
SOIL	16	41400	19.70%	3.15%	
		210170	100.00%	51.43%	OVERALL PERF
	82(PIF)				
URB	94	20852	9.88%	9.29%	
WAT	94	1868	0.89%	0.83%	
RES	76	59285	28.09%	21.35%	
VEG	78	108459	51.40%	40.09%	
SOIL	52	20565	9.75%	5.07%	
		211029	100.00%	76.63%	OVERALL PERF